APUNTES

ECONOMETRÍA

Unión de Estudiantes de Ciencias Económicas | AECUC3M

Tema 1 Modelización y Datos Econométricos

<table>
<tr><th colspan="2">Concepto de Modelo Econométrico</th></tr>
<tr><td>Concepto de Econometría</td><td>

La Econometría se basa en el desarrollo de métodos estadísticos destinados a estimar las relaciones económicas.
La Econometría se centra en los problemas inherentes a la recopilación y al análisis de datos económicos no experimentales.
La Econometría es una rama de la Estadística basada en el planteamiento de modelos que relacionen variables económicas.
Uno de los objetivos principales de la Econometría es realizar análisis causal, que permite determinar los efectos de ciertas políticas; y caracterizar y cuantificar la relación de comportamiento entre variables económicas.

</td></tr>
<tr><td>Componentes del modelo econométrico</td><td>

Las variables:

Variables observacionales: Son aquellas que se pueden medir.
<ol type="a">
Variable Dependiente: Es aquella que tratamos de explicar.
Variable Independiente: Es aquella que explica o causa a la variable dependiente. Dicho de otra forma, la variable dependiente estará causada por las variables independientes.

<table>
<tr><th>y</th><th>x</th></tr>
<tr><td>Variable dependiente</td><td>Variable independiente</td></tr>
<tr><td>Variable explicada</td><td>Variable explicativa</td></tr>
<tr><td>Variable respuesta</td><td>Variable de control</td></tr>
<tr><td>Variable predicha</td><td>Variable predictor</td></tr>
<tr><td>Regresando</td><td>Regresor</td></tr>
</table>

Variables no observacionales: Son aquellas que no se pueden medir por estar contenidas en el término de error.

Los Parámetros: Si están contenidos en la ecuación del modelo, se llamarán parámetros estructurales.
Los Datos:

Datos de Sección Cruzada o Corte Transversal:
<ol type="a">
Definición: Son conjuntos de observaciones de una o más variables medidas en distintas unidades económicas.
Ejemplo: Salario de un grupo de trabajadores de una empresa; Ventas de un conjunto de empresas en 2009, etc.

Datos de Serie Temporal:
<ol type="a">
Definición: Son conjuntos de observaciones de una o más variables medidas a lo largo del tiempo en períodos regulares.
Ejemplos: El IPC mensual, el crecimiento del PIB trimestral,etc.

Datos de Panel:

</td></tr>
</table>

	a) Definición: Son conjunto de observaciones de una o más variables medidas en distintas unidades económicas, en dos períodos distintos de tiempo, entre los que se produce un cambio estructural. b) Ejemplo: Medimos el precio de un conjunto de viviendas antes y después de construir un verdadero próximo.
Propiedades de las Esperanzas, Covarianzas ,Varianzas, Sumatorios y Límites Probabilísticos	
Propiedades de las Esperanzas	• $E(k) = k$ • $E(k \pm X) = k \pm E(X)$ • $E(k * X) = k * E(X)$ • $E(X \pm Y \pm Z) = E(X) \pm E(Y) \pm E(Z)$ • $E(X * Y) = E(X) * E(Y), si\ X\ e\ Y\ están\ incorrelacionadas\ o\ son\ independientes.$
Propiedades de las Varianzas	• $Var(k) = 0$ • $Var(k \pm X) = Var(X)$ • $Var(k * X) = k^2 * Var(X)$ • $Var(X \pm Y) = Var(X) + Var(Y) \pm 2 * Cov(X,Y)$ • $Var(X \pm Y) = Var(X) + Var(Y), si\ X\ e\ Y\ están\ incorrelacionadas\ o\ son\ independientes.$ • $Var(aX \pm bY) = Var(aX) + Var(bY) \pm 2Cov(aX, bY) = a^2Var(X) + b^2Var(Y) \pm 2 * a * b * Cov(X,Y)$
Propiedades de las Covarianzas	• $Cov(k, X) = 0$ • $Cov(X, X) = Var(X)$ • $Cov(X,Y) = Cov(Y,X)$ • $Cov(aX, bY) = a * b * Cov(X,Y)$ • $Cov(X \pm Y, Z) = Cov(X,Z) \pm Cov(Y,Z)$ • $Cov(a + bX + CY, Z) = Cov(a, Z) + Cov(bX, Z) + Cov(CY, Z) = b * Cov(X,Z) + c * Cov(Y,Z)$
Propiedades de los Sumatorios	• La suma del producto de una constante por una variable, es igual a k veces la sumatoria de la variable. $\displaystyle\sum_{i=1}^{n} kX_i = k\sum_{i=1}^{n} X_i$ • El sumatorio hasta N de una constante, es igual a N veces la constante. $\displaystyle\sum_{i=1}^{n} k = n * k$ • El sumatorio de una suma es igual a la suma de las sumatorias de cada término. $\displaystyle\sum_{i=1}^{n} (X_i + Y_i) = \sum_{i=1}^{n} X_i + \sum_{i=1}^{n} Y_i$ • El sumatorio de un producto **no es igual** al producto de las sumatorias de cada término. $\displaystyle\sum_{i=1}^{n} X_i * Y_i \neq \sum_{i=1}^{n} X_i * \sum_{i=1}^{n} Y_i$

	• El sumatorio de los cuadrados de los valores de una variable **no es igual** a la sumatoria de la variable elevado al cuadrado. $\displaystyle\sum_{i=1}^{n} X_i^2 \neq \left(\sum_{i=1}^{n} X_i\right)^2$ • Aplicación de las propiedades de los Sumatorios: 1. Cálculo de la Varianza de X: $\displaystyle\sum_{i=1}^{n}(X_i-\bar{X})^2 = \sum_{i=1}^{n}(X_i^2 - 2X_i\bar{X} + \bar{X}^2) = \sum_{i=1}^{n} X_i^2 - 2\sum_{i=1}^{n} X_i\bar{X} + \sum_{i=1}^{n}\bar{X}^2$ $\displaystyle= \sum_{i=1}^{n} X_i^2 - 2\sum_{i=1}^{n} X_i\bar{X} + n * \frac{\sum_{i=1}^{n} X_i}{n} * \bar{X} = \sum_{i=1}^{n} X_i^2 - 2\sum_{i=1}^{n} X_i\bar{X} + \sum_{i=1}^{n} X_i\bar{X}$ $\displaystyle= \sum_{i=1}^{n} X_i^2 - \sum_{i=1}^{n} X_i\bar{X} = \sum_{i=1}^{n} X_i(X_i-\bar{X})$ 2. Cálculo de la Covarianza entre X e Y: $\displaystyle\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y}) = \sum_{i=1}^{n}(X_iY_i - \bar{X}Y_i - \bar{Y}X_i + \bar{X}\bar{Y}) = \sum_{i=1}^{n} X_iY_i - \sum_{i=1}^{n}\bar{X}Y_i - \sum_{i=1}^{n}\bar{Y}X_i + \sum_{i=1}^{n}\bar{X}\bar{Y}$ $\displaystyle= \sum_{i=1}^{n} X_iY_i - \bar{X}\sum_{i=1}^{n} Y_i - \bar{Y}\sum_{i=1}^{n} X_i + n * \frac{\sum_{i=1}^{n} X_i}{n} * \bar{Y}$ $\displaystyle= \sum_{i=1}^{n} X_iY_i - \bar{X}\sum_{i=1}^{n} Y_i - \bar{Y}\sum_{i=1}^{n} X_i + \bar{Y}\sum_{i=1}^{n} X_i = \sum_{i=1}^{n} X_iY_i - \bar{X}\sum_{i=1}^{n} Y_i$ $\displaystyle= \sum_{i=1}^{n} Y_i(X_i-\bar{X})$
Propiedades de Límites Probabilísticos	• El límite probabilístico de un constante es igual al constante. $\displaystyle p\lim_{n\to\infty}(k) = k$ • El límite probabilístico del producto de un constante por una variable es igual al constante por el límite probabilífico de la variable. $\displaystyle p\lim_{n\to\infty}(kX) = k * p\lim_{n\to\infty}(X)$ • El límite probabilístico de una suma o resta de dos variables es igual a la suma o resta de límites probabilísticos de las dos variables. $\displaystyle p\lim_{n\to\infty}(X \pm Y) = p\lim_{n\to\infty}(X) \pm p\lim_{n\to\infty}(Y)$ • El límite probabilístico de un producto entre dos variables es igual al producto de los límites probabilísticos de las dos variables.

	$$p\lim_{n\to\infty}(X*Y) = p\lim_{n\to\infty}(X) * p\lim_{n\to\infty}(Y)$$ • El límite probabilístico de un cociente entre dos variables es igual al cociente de los límites probabilísticos de las dos variables. $$p\lim_{n\to\infty}\left(\frac{X}{Y}\right) = \frac{p\lim_{n\to\infty}(X)}{p\lim_{n\to\infty}(Y)}$$
Concepto de Esperanza Condicionada y Causalidad CETERIS PARIBUS	
Introducción	• El objetivo de los modelos econométricos es predecir Y en función de X. • Dando por hecho que entr ambas variables existe una relación de causalidad, es decir, que Y está causada por X puede demostrarse que la mejor predicción de Y en función de X es la Esperanza de Y condicionada a X, E(Y/X), o Esperanza Condicional. También se le conoce como Función de Regresión Poblacional.
Efecto Causal	• El objetivo de los economistas es inferir si una variable tiene un efecto causal sobre otras variables. • Para obtener este efecto causal, debemos aislar el efecto de las otras variables influyentes en el salario. • Esto es lo que se conoce como causalidad CETERIS PARIBUS, es decir, mantenemos constantes el resto de variables influyentes en Y.
Propiedades de las Esperanzas Condicionadas	• $E(k/X) = k$ • Suponemos que Z=g(X) $E(Z/X) = E(g(X)/X) = g(X)$ • Suponemos que Z=a+bX+cY $E(Z/X) = E(a + bX + cY/X) = E(a/X) + E(bX/X) + E(cY/X) = a + bX + c * E(Y/X)$ • Ley de Esperanzas Iteradas: $E(Y) = E[E(Y/X)]$ • Descomposición de la Varianza Marginal: $Var(Y) = E[Var(Y/X)] + Var[E(Y/X)]$

Tema 2 Modelo de Regresión Lineal Simple

Concepto y Tipos de Predicción	
Concepto	• El objetivo del Modelo de Regresión Lineal Simple es predecir la variable dependiente Y, utilizando la información de la variable independiente X. • El concepto de predicción consiste en elaborar una predicción c(X), al ser una función de x. Evidentemente cometeremos un error de predicción (u), que será la diferencia entre el valor real de Y y nuestra predicción c(X): $$U = Y - c(X)$$ • Como el error de predicción (u) puede ser positivo, negativo o nulo, para asegurarnos de que siempre obtengamos magnitudes positivas, definimos el Error Cuadrático Medio de la Predicción ECM (u): $$ECM(U) = E[Y - c(X)]^2$$ • Nuestro objetivo es encontrar la predicción c(X) que haga que el ECM (u) sea mínimo.
Tipos de Predicción	• Mejor Predicción Constante: a) Supuesto: No conocemos la información de la variable X y la función c(X)=c $$ECM(U) = E[Y - C]^2$$ b) Minimizando esta función, obtenemos la mejor predicción constante de la variable Y. c) Predictor: La media poblacional es el mejor predictor constante de la variable Y. $$c = E(Y)$$ • Mejor Predicción Lineal: a) Supuesto: Conocemos la variable X para predecir la variable Y, y la función c(X) es de tipo lineal, de tal manera que $c(X) = c_0 + c_1 * X$. $$ECM(U) = E[Y - c(X)]^2 = E[Y - (c_0 + c_1 * X)]^2 = E[Y - c_0 - c_1 * X]^2$$ b) Minimizando este error cuadrático medio, obtenemos la mejor predicción lineal, y la llamamos Proyección Lineal de Y/X o Predictor Lineal Óptimo de Y/X. $$L(Y/X) = \alpha_0 + \alpha_1 * X$$ $$\alpha_0 = E(Y) - \alpha_1 E(X)$$ $$\alpha_1 = \frac{C(X,Y)}{V(X)}$$ $$L(Y/X) = \alpha_0 + \alpha_1 * X = E(Y) + \frac{C(X,Y)}{V(X)}(X - E(X))$$ • La Mejor Predicción: a) Supuesto: Conocemos la variable X antes de hacer la predicción de la variable Y. En este caso, la función c(X) puede ser lineal o no lineal, pero en general es desconocida. $$ECM(U) = E[Y - c(X)]^2$$ b) Minimizando esta función, obtenemos la mejor predicción de la variable Y, que será la Esperanza de Y condicionada a X, E(Y/X), o Función de Esperanza Condicional. c) Si E(Y/X) es Lineal, entonces coincidirá con la proyección lineal: $$E(Y/X) = L(Y/X) = \alpha_0 + \alpha_1 * X$$

<table>
<tr><td></td><td>d) Si E(Y/X) no es lineal, entonces no coincidirá con la proyección lineal.
$$E(Y/X)=\alpha_0+\alpha_1*X+\alpha_2*X^2\neq L(Y/X)=\alpha_0+\alpha_1*X$$</td></tr>
<tr><th colspan="2">Supuestos del Modelo de Regresión Lineal Simple</th></tr>
<tr><td>Linealidad en los Parámetros del Modelo</td><td>$$Y=\beta_0+\beta_1*X+\varepsilon$$
• Esto quiere decir, que el modelo será lineal si los parámetros (β_0 , β_1) son lineales.
• Por lo tanto, un modelo puede no ser lineal en variables, pero sólo es en perámetros, se considerará que cumple este supuesto.</td></tr>
<tr><td>$\boldsymbol{E(\varepsilon/X)=0}$</td><td>• Suponemos que $\boldsymbol{E(\varepsilon/X)=0}$, donde ε es el término de error o peturbación, y es una variable aleatoria que recoge el efecto de las variables explicativas no incluidas en el modelo.
• Implicaciones del Supuesto:
1. Se cumple que $\boldsymbol{E(\varepsilon)=0}$.
2. La $\boldsymbol{Cov(X,\varepsilon)=0}$, Incorrelación entre X y ε, también se puede decir que X es exógeno.
3. La función de Esperanza condicional será lineal, es decir, $\boldsymbol{E(Y/X)=\beta_0+\beta_1*X}$</td></tr>
<tr><td>Homocedasticidad</td><td>• Supone que la Varianza Condicionada del Error es Constante e igual a σ^2.
$$\boldsymbol{V(\varepsilon/X)=E\left(\varepsilon^2/X\right)=\sigma^2}$$
• Implicaciones del Supuesto:
1. $V(\varepsilon)=\sigma^2$
2. $V(Y/X)=\sigma^2$</td></tr>
<tr><th colspan="2">Estimación del Modelo: Mínimos Cuadrados Ordinarios</th></tr>
<tr><td>Principio de Analogía</td><td>• Las expresiones que se cumplen en términos muestrales también se cumplirán en términos poblacionales (y viceversa).
<table>
<tr><th>POBLACIONAL</th><th>MUESTRAL</th></tr>
<tr><td>$\beta_0=E(Y)-\alpha_1E(X)$</td><td>$\widehat{\beta_0}=\bar{y}-\widehat{\beta_1}*\bar{x}$</td></tr>
<tr><td>$\beta_1=\dfrac{C(X,Y)}{V(X)}$</td><td>$\widehat{\beta_1}=\dfrac{s_{xy}}{s_x^2}$</td></tr>
<tr><td>$E(Y/X)=\beta_0+\beta_1*X$</td><td>$\hat{y}=\widehat{\beta_0}+\widehat{\beta_1}*x$</td></tr>
</table></td></tr>
<tr><td>Condiciones de Primer Orden para las Estimaciones MCO</td><td>• Las condiciones de primer orden para las estimaciones MCO es una expresión que viene del cálculo de optimización.
• La expresión Mínimos Cuadrados Ordinarios viene del hecho de que estos valores estimados minimizan la suma de los cuadrados de los residuos.
• Minimizar la Suma de los Cuadrados de los Residuos:
Dado que $\hat{y}=\widehat{\beta_0}+\widehat{\beta_1}*x+\hat{\varepsilon}$, por lo que $\hat{\varepsilon}=\hat{y}-\widehat{\beta_0}-\widehat{\beta_1}*x$
Nuestro objetivo es minimizar la suma de los residuos al cuadrado, es decir:
$$\min_{\widehat{\beta_0},\widehat{\beta_1}}\sum_{i=1}^{n}\left(\hat{y}-\widehat{\beta_0}-\widehat{\beta_1}*x\right)^2$$
Una condición necesaria para que $\widehat{\beta_0}$ y $\widehat{\beta_1}$ sean soluciones del problema de minimización es</td></tr>
</table>

	que las dedivadas parciales de la función sea igual a 0. $$\frac{\partial f}{\partial \widehat{\beta_0}} = 0;\ -2\sum_{i=1}^{n}\left(\hat{y} - \widehat{\beta_0} - \widehat{\beta_1} * x\right) = 0; \sum_{i=1}^{n}\left(\hat{y} - \widehat{\beta_0} - \widehat{\beta_1} * x\right) = 0;$$ $$\frac{\partial f}{\partial \widehat{\beta_1}} = 0;\ -2\sum_{i=1}^{n} x\left(\hat{y} - \widehat{\beta_0} - \widehat{\beta_1} * x\right) = 0; \sum_{i=1}^{n} x\left(\hat{y} - \widehat{\beta_0} - \widehat{\beta_1} * x\right) = 0$$ Por lo tanto observamos que para que $\widehat{\beta_0}$ sea un estimador óptimo, la condición es que $\sum_{i=1}^{n}(\varepsilon) = 0$, y para que $\widehat{\beta_1}$ sea un estimador óptimo, la condición es que $\sum_{i=1}^{n} x\varepsilon = 0$. Y estas dos condiciones son las dos implicaciones de uno de los supuestos del MRLS: $\boldsymbol{E(\varepsilon) = 0}$ y $\boldsymbol{Cov(X,\varepsilon) = E(X\varepsilon) = 0}$.
Propiedades Algebraicas de los Estimadores MCO	• La suma, y por lo tanto, la media muestral de los residuos MCO es nula: $$\sum_{i=1}^{n}(\varepsilon) = 0$$ Esta propiedad deriva directamente de la condición de primer orden de los MCO. • La Covarianza muestral entre los regresores y los residuos MCO es nula: $$\sum_{i=1}^{n} x\varepsilon = 0$$ • El punto $(\bar{x}, \bar{y})$ siempre está sobre la recta de regresión MCO, es decir $$\bar{y} = \widehat{\beta_0} + \widehat{\beta_1} * \bar{x}$$ • Aplicando las propiedades anteriores, observamos que $y = \hat{y} + \varepsilon$, y dado que $\boldsymbol{E(\varepsilon) = 0}$, por lo tanto, $\bar{y} = \bar{\hat{y}}$.
Estimación de la varianza del error	• Distinción entre los errores y los residuos: Los errores **nunca son observables**, y aparecen en la ecuación que contiene los parámetros poblacionales: $$Y = \beta_0 + \beta_1 * X + \varepsilon\ (error)$$ Los residuos se calculan **a partir de los datos**, y aparecen en la ecuación estimada: $$\hat{\varepsilon}\ (residuo) = y - \widehat{\beta_0} - \widehat{\beta_1} * X = (\beta_0 + \beta_1 * X + \varepsilon) - \widehat{\beta_0} - \widehat{\beta_1} * X = \varepsilon - \left(\widehat{\beta_0} - \beta_0\right) - \left(\widehat{\beta_1} - \beta_1\right) * X$$ • Estimación de la varianza del error: $$\sigma^2 = E(\varepsilon^2)$$ Luego, un estimador de σ^2 será: $$\widetilde{\sigma^2} = \frac{\sum_{i=1}^{n}\hat{\varepsilon}^2}{n} = \frac{SCE}{n}$$

	Sin embargo, este estimador tiene un sesgo, porque no considera las restricciones que los residuos MCO deben satisfacer. Estas restricciones vienen dadas por las dos condiciones de primer orden del estimador MCO: $\sum_{i=1}^{n}(\varepsilon) = 0$ y $\sum_{i=1}^{n} x\varepsilon = 0$. Un estimador insesgado de σ^2 será: $$\widehat{\sigma^2} = \frac{\sum_{i=1}^{n} \hat{\varepsilon}^2}{n-2} = \frac{SCE}{n-2}$$
Propiedades de los Estimadores MCO	
Teorema de Gauss-Markov	• De entre todos los estimadores, los estimadores MCO son lineales, insesgados y tienen varianza mínima, es decir, que son óptimos.
Linealidad	• Significa que $\widehat{\beta_0}$ y $\widehat{\beta_1}$ son en realidad, funciones lineales de Y. $$\widehat{\beta_0} = \bar{y} - \widehat{\beta_1} * \bar{x}$$ $$\widehat{\beta_1} = \frac{s_{xy}}{s_x^2} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \sum_{i=1}^{n} c_i Y_i \, , siendo \; c_i = \frac{(x_i - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \, , y \; Y_i = y_i - \bar{y}$$
Insesgadez	• Decimos que un estimador es insesgado o centrado cuando su esperanza es igual al parámetro que estimamos. $$E(\widehat{\beta_0}) = \beta_0$$ $$E(\widehat{\beta_1}) = \beta_1$$
Varianza mínima	• Puede demostrarse que $\widehat{\beta_0}$ y $\widehat{\beta_1}$ de MCO tienen mínima varianza, y por tanto, serán los estimadores más eficientes. $$V(\widehat{\beta_0}) = \frac{\sigma^2}{n * s_x^2}$$ $$V(\widehat{\beta_1}) = \frac{\sigma^2 * \sum_{i=1}^{n} x_i^2}{n * s_x^2}$$
Consistencia	• Los estimadores $\widehat{\beta_0}$ y $\widehat{\beta_1}$ son estimadores consistentes si a medida que se incrementa el tamaño muestral, su valor se aproxima al verdadero parámetro poblacional. En términos econométricos, la consistencia equivale a convergencia en probabilidad. $$p \lim_{n \to \infty} \widehat{\beta_J} = \beta_j$$
Interpretación de los coeficientes	
Introducción	• Al plantear cualquier modelo de regresión, nuestro objetivo es interpretar el efecto de x sobre y. • Para que los parámetros de un modelo puedan interpretarse, el modelo tiene que ser lineal en parámetros. • Como vimos, la condición para que el modelo sea lineal era el cumplimiento del $\boldsymbol{E(\varepsilon/X) = 0}$. Esto viene a significar que, si $\boldsymbol{E(\varepsilon/X) \neq 0}$,entonces los parámetros no tendrán interpretación. • Concepto de Variación Absoluta, Variación Relativa y Variación Porcentual: Ejemplo: Supongamos que el salario de un individuo (x) al mes son 2000€, y su salario se incrementa hasta 2500€. Vamos a calcular la variación absoluta, relativa y porcentual del salario: Variación Absoluta: $$\Delta x = x_1 - x_0 = 2500 - 2000 = 500€$$

	Variación Relativa: $$\frac{\Delta x}{x}=\frac{x_1-x_0}{x_0}=\frac{2500-2000}{2000}=\frac{500}{2000}=0{,}25$$ Variación Porcentual: $$\frac{\Delta x}{x}*100=0{,}25*100=25\%$$
Concepto de Elasticidad y Semielasticidad	• **Elasticidad**: Cuando las 2 variables están expresados en función de Ln. $$Ln(Y)=\beta_0+\beta_1*Ln(X)$$ • **Semielasticidad**: Cuando sólo 1 de las variables se ve afectado por Ln. $$Y=\beta_0+\beta_1*Ln(X)$$
Modelo Lineal	$$Y=\beta_0+\beta_1*X+\varepsilon$$ • Interpretación de β_1: $$\beta_1=\frac{\Delta Y}{\Delta X}\ ;\ \Delta\boldsymbol{Y}=\boldsymbol{\beta_1}*\Delta\boldsymbol{X}$$ Ante un incremento de 1 unidad en X, le corresponde en promedio una variación de β_1 unidades en Y. • Elasticidad de Y con respecto a X: $$E_{Y/X}=\frac{\frac{\Delta Y}{Y}}{\frac{\Delta X}{X}}=\frac{\Delta Y}{\Delta X}*\frac{X}{Y}=\beta_1*\frac{X}{Y}$$ En este modelo, la elasticidad no es constante, puesto que depende de los valores de X e Y.
Modelo Logarítmico-Lineal	$$Ln(Y)=\beta_0+\beta_1*X+\varepsilon$$ • Interpretación de β_1: $$\beta_1=\frac{\Delta Ln(Y)}{\Delta X}\approx\frac{\frac{\Delta Y}{Y}}{\Delta X}=\frac{\Delta Y}{\Delta X}*\frac{1}{Y}\ ;\ 100*\beta_1=\frac{\frac{\Delta Y}{Y}}{\Delta X}*100\ ;\ \%\Delta\boldsymbol{Y}=(\boldsymbol{100}*\boldsymbol{\beta_1})\Delta\boldsymbol{X}$$ Ante un incremento de 1 unidad en X, le corresponde en promedio una variación de $100*\beta_1$ puntos porcentuales en Y. • Elasticidad de Y con respecto a X: $$E_{Y/X}=\frac{\frac{\Delta Y}{Y}}{\frac{\Delta X}{X}}=\frac{\Delta Y}{\Delta X}*\frac{X}{Y}=\beta_1*X$$ En este modelo, la elasticidad no es constante, puesto que depende del valor de X.
Modelo Lineal-Logarítmico	$$Y=\beta_0+\beta_1*Ln(X)+\varepsilon$$ • Interpretación de β_1: $$\beta_1=\frac{\Delta Y}{\Delta Ln(X)}\approx\frac{\Delta Y}{\frac{\Delta X}{X}}=\frac{\Delta Y}{\Delta X}*X\ ;\ 100*\beta_1=\frac{\Delta Y}{\frac{\Delta X}{X}}*100\ ;\ \Delta\boldsymbol{Y}=\left(\frac{\boldsymbol{\beta_1}}{\boldsymbol{100}}\right)\%\Delta\boldsymbol{X}$$ Ante un incremento de 1 punto porcentual en X, le corresponde en promedio una variación de $\frac{\beta_1}{100}$ unidades en Y. • Elasticidad de Y con respecto a X:

	$$E_{Y/X} = \frac{\frac{\Delta Y}{Y}}{\frac{\Delta X}{X}} = \frac{\Delta Y}{\Delta X} * \frac{X}{Y} = \beta_1 * \frac{1}{Y}$$ En este modelo, la elasticidad no es constante, puesto que depende del valor de Y.
Modelo Doble Logarítmico	$$Ln(Y) = \beta_0 + \beta_1 * Ln(X) + \varepsilon$$ • Interpretación de β_1: $$\beta_1 = \frac{\Delta Ln(Y)}{\Delta Ln(X)} \approx \frac{\frac{\Delta Y}{Y}}{\frac{\Delta X}{X}} = \frac{\Delta Y}{\Delta X} * \frac{X}{Y}\ ;\ 100 * \beta_1 = \frac{\frac{\Delta Y}{Y}}{\frac{\Delta X}{X}} * 100\ ;\ \%\Delta \boldsymbol{Y} = \boldsymbol{\beta_1} \%\Delta \boldsymbol{X}$$ Ante un incremento de 1 punto porcentual en X, le corresponde en promedio una variación de β_1 puntos porcentuales en Y. • Elasticidad de Y con respecto a X: $$E_{Y/X} = \frac{\frac{\Delta Y}{Y}}{\frac{\Delta X}{X}} = \frac{\Delta Y}{\Delta X} * \frac{X}{Y} = \beta_1$$ En este modelo, la elasticidad es constante.
Unidades de Medida	• Los estimadores MCO cambian en una forma totalmente previsible cuando se modifican las unidades de medida de las variables dependientes e independientes. • Si la variable dependiente (Y) se multiplica por la constante c, entonces los valores estimados MCO del término constante y de la pendiente también se multiplican por c. $$c * Y = c * (\beta_0 + \beta_1 * X + \varepsilon) = c * \beta_0 + c * \beta_1 * X + c * \varepsilon$$ • Si la variable independiente (X) se divide o se multiplica por una constante no nula, c , entonces el coeficiente de la pendiente MCO se multiplica o se divide por c, respectivamente. El cambio sólo en las unidades de medida de la variable independiente no afecta al término constante. $$Y = \beta_0 + c * \beta_1 * X + \varepsilon$$
Bondad de Ajuste	
Suma de Cuadrados	• Suma de Cuadrados Total (SCT): SCT=SCM+SCE $$\sum_{i=1}^{n} (y_i - \bar{y})^2$$ • Suma de Cuadrados del Modelo (SCM): $$\sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2$$ • Suma de Cuadrados de los residuos (SCE): $$\sum_{i=1}^{n} \hat{\varepsilon}^2$$
R-cuadrado	• Es una forma de medir la capacidad de la variable independiente o explicativa X de explicar la variable dependiente. • **R-cuadrado** es la proporción de la variación explicada en comparación con la variación total.

	• Se interpreta como la fracción de la variación muestral en Y que viene explicada por X. • $0 \leq R^2 \leq 1$ • 100*R^2es el porcentaje de la variación muestral de Y que viene explicada por X. $$R^2 = \frac{SCM}{SCT} = 1 - \frac{SCE}{SCT}$$ • R^2 también es igual al cuadrado del coeficiente de correlación muestral entre y y $\hat{y_i}$. $$R^2 = \widehat{\rho_{y\hat{y}_i}}^2 = \frac{S_{y\hat{y}_i}}{S_y S_{\hat{y}_i}}$$

Tema 3 Modelo de Regresión Múltiple

Modelo con k variables explicativas	
Supuestos del Modelo de Regresión Múltiple	• **Linealidad en los parámetros**: $$Y = \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \cdots + \beta_k * X_k + \varepsilon$$ Esto quiere decir, que el modelo será lineal si los parámetros ($\beta_0, \beta_1, \beta_2 \dots \beta_k$) son lineales. Por lo tanto, un modelo puede no ser lineal en variables, pero sólo es en perámetros, se considerará que cumple este supuesto. • $\boldsymbol{E(\varepsilon/X_1, X_2, \dots, X_k) = 0}$: Se trata de un supuesto fundamental del modelo y tiene las siguientes implicaciones: 4. Se cumple que $\boldsymbol{E(\varepsilon) = 0}$. 5. La $\boldsymbol{Cov(X_1, X_2, \dots, X_k; \varepsilon) = 0}$, Incorrelación entre $X_1, X_2, \dots, X_k$ y ε. 6. La función de Esperanza condicional será lineal, es decir: $$\boldsymbol{E(Y/X_1, X_2, \dots, X_k) = \beta_0 + \beta_1 * X + \beta_2 * X_2 + \cdots + \beta_k * X_k}$$ • **Homocedasticidad**: Supone que la Varianza Condicionada del Error es Constante e igual a σ^2. $$\boldsymbol{V(\varepsilon/X_1, X_2, \dots, X_k) = E(\varepsilon^2/X_1, X_2, \dots, X_k) = \sigma^2}$$ Implicaciones del Supuesto: 3. $V(\varepsilon) = \sigma^2$ 4. $V(Y/X_1, X_2, \dots, X_k) = \sigma^2$ • **No Colinealidad Perfecta**: Supone que no existe ninguna combinación lineal exacta entre variables explicativas, porque si existiese, el modelo presentará multicolinealidad o Colinealidad y no podrá estimarse. Es importante tener en cuenta que el supuesto sí permite que las variables independientes estén correlacionadas, lo que no pueden es estar perfectamente correlacionadas. Ejemplos de combinaciones exactas de variables: $$X_1 + X_2 = 1$$ $$X_1 - 2X_2 = X_3$$ $$X_2 = 2X_1$$
Modelo con dos variables explicativas	$$Y = \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \varepsilon$$ $$\beta_0 = E(Y) - \beta_1 E(X_1) - \beta_2 E(X_2)$$ $$\beta_1 = \frac{V(X_2) * C(X_1, Y) - C(X_1, X_2) * C(X_2, Y)}{V(X_1) * V(X_2) - [C(X_1, X_2)]^2}$$ $$\beta_2 = \frac{V(X_1) * C(X_2, Y) - C(X_1, X_2) * C(X_1, Y)}{V(X_1) * V(X_2) - [C(X_1, X_2)]^2}$$ Un caso especialmente importante en este modelo es el caso de Incorrelación entre X_1 y X_2, es decir, $C(X_1, X_2) = 0$. Por lo tanto:

	$\beta_1 = \frac{V(X_2)*C(X_1,Y)}{V(X_1)*V(X_2)} = \frac{C(X_1,Y)}{V(X_1)}$ $\beta_2 = \frac{V(X_1)*C(X_2,Y)}{V(X_1)*V(X_2)} = \frac{C(X_2,Y)}{V(X_2)}$
Estimación del Modelo de Regresión Múltiple: MCO	
Condición de Primer Orden	Dado que $\hat{y} = \widehat{\beta_0} + \widehat{\beta_1} * X_1 + \widehat{\beta_2} * X_2 + \cdots + \widehat{\beta_k} * X_k + \hat{\varepsilon}$, por lo que $\hat{\varepsilon} = \hat{y} - \widehat{\beta_0} - \widehat{\beta_1} * X_1 - \widehat{\beta_2} * X_2 - \cdots - \widehat{\beta_k} * X_k$ Nuestro objetivo es minimizar la suma de los residuos al cuadrado, es decir: $\min_{\beta_0,\beta_1} \sum_{i=1}^{n} (\hat{y} - \widehat{\beta_0} - \widehat{\beta_1} * X_1 - \widehat{\beta_2} * X_2 - \cdots - \widehat{\beta_k} * X_k)^2$ Una condición necesaria para que $\widehat{\beta_0}$ y $\widehat{\beta_1}$ sean soluciones del problema de minimización es que las dedivadas parciales de la función sea igual a 0. $\frac{\partial f}{\partial \widehat{\beta_0}} = 0;\ -2\sum_{i=1}^{n} (\hat{y} - \widehat{\beta_0} - \widehat{\beta_1} * X_1 - \widehat{\beta_2} * X_2 - \cdots - \widehat{\beta_k} * X_k) = 0;$ $\sum_{i=1}^{n} (\hat{y} - \widehat{\beta_0} - \widehat{\beta_1} * X_1 - \widehat{\beta_2} * X_2 - \cdots - \widehat{\beta_k} * X_k) = 0;$ $\frac{\partial f}{\partial \widehat{\beta_1}} = 0;\ -2\sum_{i=1}^{n} X_1(\hat{y} - \widehat{\beta_0} - \widehat{\beta_1} * X_1 - \widehat{\beta_2} * X_2 - \cdots - \widehat{\beta_k} * X_k) = 0;$ $\sum_{i=1}^{n} X_1(\hat{y} - \widehat{\beta_0} - \widehat{\beta_1} * X_1 - \widehat{\beta_2} * X_2 - \cdots - \widehat{\beta_k} * X_k) = 0$ $\frac{\partial f}{\partial \widehat{\beta_2}} = 0;\ -2\sum_{i=1}^{n} X_2(\hat{y} - \widehat{\beta_0} - \widehat{\beta_1} * X_1 - \widehat{\beta_2} * X_2 - \cdots - \widehat{\beta_k} * X_k) = 0;$ $\sum_{i=1}^{n} X_2(\hat{y} - \widehat{\beta_0} - \widehat{\beta_1} * X_1 - \widehat{\beta_2} * X_2 - \cdots - \widehat{\beta_k} * X_k) = 0$
Propiedades Algebraicas de los Estimadores MCO	• La media muestral de los residuos es cero. $\sum_{i=1}^{n}(\varepsilon) = 0$ • La Covarianza muestral entre cada variable independiente y los residuos MCO es cero, de lo que se desprende que la covarianza muestral entre los valores ajustados MCO y los residuos MCO es cero. $\sum_{i=1}^{n} x_{1\ldots k}\varepsilon = 0$ • El punto$(\overline{X_1}, \overline{X_2}, \ldots, \overline{X_K}\ ; \overline{Y})$ siempre está sobre la recta de regresión MCO.
Interpretación de los Coeficientes en un Modelo de Regresión Múltiple	
Interpretar la ecuación de la regresión MCO para 2 variables	$\hat{y} = \widehat{\beta_0} + \widehat{\beta_1} * X_1 + \widehat{\beta_2} * X_2$ • El parámetro del término constante $\widehat{\beta_0}$ es el valor predicho de y cuando X_1=X_2=0. Siempre se necesita el parámetro del término constante para obtener una predicción de y a partir de la recta de regresión MCO. • Los valores estimados $\widehat{\beta_1}$ y $\widehat{\beta_2}$ tienen interpretaciones de **efecto parcial** o **ceteris paribus**.

independientes	$$\Delta\hat{y} = \widehat{\beta_1}\Delta X_1 + \widehat{\beta_2}\Delta X_2$$ Cuando X_2 se mantiene fijo, de forma que $\Delta X_2 = 0$, entonces: $$\Delta\hat{y} = \widehat{\beta_1}\Delta X_1$$ Cuando X_1 se mantiene fijo, de forma que $\Delta X_1 = 0$, entonces: $$\Delta\hat{y} = \widehat{\beta_2}\Delta X_2$$
Interpretar la ecuación de la regresión MCO para más de 2 variables independientes	$$\hat{y} = \widehat{\beta_0} + \widehat{\beta_1} * X_1 + \widehat{\beta_2} * X_2 + \cdots + \widehat{\beta_k} * X_k + \hat{\varepsilon}$$ • Escrito en términos de cambios: $$\Delta\hat{y} = \widehat{\beta_1}\Delta X_1 + \widehat{\beta_2}\Delta X_2 + \cdots + \widehat{\beta_k}\Delta X_k$$ • El coeficiente de X_1 mide el cambio en $\hat{y}$ por cada incremento en una unidad de X_1, manteniendo fijas las restantes variables independientes. $$\Delta\hat{y} = \widehat{\beta_1}\Delta X_1$$ • La utilidad del análisis de regresión múltiple reside en que nos proporciona **una interpretación ceteris paribus** aun cuando los datos no hayan sido recogidos de una forma ceteris paribus. • La utilidad del análisis de regresión múltiple reside en que nos permite hacer **en un medio no experimental** lo que los cientficos hacen en el medio controlado de un laboratorio: mantener fijos el resto de los factores.
Formas Cuadráticas	• Cuando planteamos el modelo de regresión, el primer supuesto que realizamos es la linealidad en los parámetros. Sin embargo, las variables no tienen porqué ser lineales. • Puede ocurrir que al plantear un modelo de regresión para recoger el efecto de alguna variable explicativa, no baste con incluirla en forma lineal, sino que tengamos que incluir un término cuadrático. $$\hat{y} = \widehat{\beta_0} + \widehat{\beta_1} * X_1 + \widehat{\beta_2} * X_2 + \widehat{\beta_3} * {X_2}^2 + \hat{\varepsilon}$$ • Cuando la variable contiene un término cuadrático, la interpretación se realiza a partir de la primera derivada del modelo: $$\frac{\partial Y}{\partial X_2} = \widehat{\beta_2} + 2\widehat{\beta_3}X_2$$ • Sustituyendo las derivadas por los incrementos: $$\frac{\Delta Y}{\Delta X_2} = \widehat{\beta_2} + 2\widehat{\beta_3}X_2$$ $$\Delta\hat{y} = (\widehat{\beta_2} + 2\widehat{\beta_3}X_2)\Delta X_2$$ • Interpretación del efecto de la variable independiente X_2: Ante un incremento de 1 unidad en X_2, le corresponde una variación en promedio de $(\widehat{\beta_2} + 2\widehat{\beta_3}X_2)$ unidades en Y, manteniendo el restos de variables independientes fijos. Por tanto, este efecto no es constante, y dependerá del valor que tome X_2.
Términos de Interacción	• Puede ocurrir que al plantear un modelo de regresión, el efecto de una variable explicativa dependa de otras variables explicativas del modelo. Cuando esto ocurre, es necesario introducir términos de interacción. $$\hat{y} = \widehat{\beta_0} + \widehat{\beta_1} * X_1 + \widehat{\beta_2} * X_2 + \widehat{\beta_3} * X_1 * X_2 + \hat{\varepsilon}$$ • Para obtener en este caso el efecto parcial ceteris paribus de X_1, habrá que plantear la siguiente derivada:

	$$\frac{\partial Y}{\partial X_1} = \widehat{\beta_1} + \widehat{\beta_3} X_2$$ • Sustituyendo las derivadas por los incrementos: $$\frac{\Delta Y}{\Delta X_1} = \widehat{\beta_1} + \widehat{\beta_3} X_2$$ $$\Delta \hat{y} = (\widehat{\beta_1} + \widehat{\beta_3} X_2)\Delta X_1$$ • Interpretación del efecto de la variable independiente X_1: Ante un incremento de 1 unidad en X_2, le corresponde una variación en promedio de $(\widehat{\beta_1} + \widehat{\beta_3} X_2)$ unidades en Y, manteniendo el restos de variables independientes fijos. Por tanto, este efecto no es constante, y dependerá del valor que tome X_2.
Propiedades de los Estimadores de MCO: Teorema de Gauss-Markow	
Teorema de Gauss-Markov	• De entre todos los estimadores, los estimadores MCO son lineales, insesgados y tienen varianza mínima, es decir, que son óptimos.
Linealidad	• Significa que $\widehat{\beta_0}, \widehat{\beta_1}, \widehat{\beta_2} \ldots, \widehat{\beta_k}$ se pueden expresar mediante una combinación lineal de Y.
Insesgadez	• Si se cumple el supuesto $\boldsymbol{E(\varepsilon / X_1, X_2, \ldots, X_k) = 0}$, entonces los estimadores MCO serán insesgados: $$E(\widehat{\beta_i}) = \beta_i$$
Varianza Mínima	• Se puede demostrar que los estimadores MCO tienen la menor varianza de entre todos los estimadores lineales e insesgados. $$V(\widehat{\beta_j}) = \frac{\sigma^2}{SCT_j * (1 - R_j^2)}$$ Donde $SCT_j = \sum_{i=1}^{n} (x_{ij} - \bar{x}_j)^2$ es la varianza muestral total de x_j, y R_j^2 es el R-cuadrado de la regresión de x_j sobre el resto de las variables independientes.
Consistencia	• Los estimadores $\widehat{\beta_0}, \widehat{\beta_1}, \widehat{\beta_2} \ldots, \widehat{\beta_k}$ son estimadores consistentes si a medida que se incrementa el tamaño muestral, su valor se aproxima al verdadero parámetro poblacional. En términos econométricos, la consistencia equivale a convergencia en probabilidad. $$p \lim_{n \to \infty} \widehat{\beta_j} = \beta_j$$
Inferencia en los Modelos de Regresión	
Contraste de Hipótesis de un único parámetro poblacional	$$Y = \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \cdots + \beta_k * X_k + \varepsilon$$ • **Hipótesis nula**: $$H_0: \beta_j = 0$$ **Interpretación**: Contrastar que x_j no tiene ningún efecto sobre el valor esperado de Y. • **Estadístico de Contraste**: $$t_{\widehat{\beta_j}} = \frac{\widehat{\beta_j}}{s(\widehat{\beta_j})}$$ 1. Como $s(\widehat{\beta_j})$ siempre es positivo, $\boldsymbol{t_{\widehat{\beta_j}}}$ **tiene el mismo signo que** $\boldsymbol{\widehat{\beta_j}}$.

	2. La estimación puntual $\widehat{\beta_J}$ nunca será exactamente igual a 0, sea o no cierta la hipótesis nula. La cuestión es saber a qué distancia está $\widehat{\beta_J}$ de 0. 3. Estamos contrastando hipótesis sobre los **parámetros poblacionales**. No estamos contrastando hipótesis sobre las estimaciones obtenidas de una muestra particular. 4. En Econometría, todas las muestras son Asintóticas (n>30), y estos estadísticos convergen a una **distribución normal**: $t_{n-k-1} \rightarrow N(0,1)$ • **Contraste frente a alternativas unilaterales**: $$H_1: \beta_j > 0$$ 1. Bajo la hipótesis alternativa $\beta_j > 0$, el valor esperado de $t_{\widehat{\beta_J}}$ es positivo. 2. Regla de Rechazo: $t_{\widehat{\beta_J}} > c$, siendo c el valor crítico. 3. Para grados de libertad **mayores que 120**, se pueden usar los valores críticos de la **normal tipificada**. $$H_1: \beta_j < 0$$ 1. Bajo la hipótesis alternativa $\beta_j < 0$, el valor esperado de $t_{\widehat{\beta_J}}$ es negativo. 2. Regla de Rechazo: $t_{\widehat{\beta_J}} < -c$, siendo c el valor crítico. • **Contraste frente a alternativas bilaterales**: $$H_1: \beta_j \neq 0$$ 1. Bajo esta alternativa, x_j tiene un efecto ceteris paribus en Y, sin especificar si el efecto es positivo o negativo. 2. Regla de Rechazo: $\left\|t_{\widehat{\beta_J}}\right\| > c$, siendo c el valor crítico. • **Interpretación del Resultado de Contraste**: 1. Si se rechaza H_0, decimos que x_j es **estadísticamente significativo**. 2. Si no se rechaza H_0, decimos que x_j es **estadísticamente no significativo**. • **Contraste de otras hipótesis sobre $\boldsymbol{\beta_j}$**: $$H_0: \beta_j = a_j$$ $$t_{\widehat{\beta_J}} = \frac{(\widehat{\beta_J} - a_j)}{s(\widehat{\beta_J})} = \frac{(valor\ estimado - valor\ hipotético)}{error\ estándar}$$ • **Cálculo de p-valores**: El p-valor es el nivel de significatividad más pequeño al que se rechazaría la hipótesis nula. $$P(\|T\| > \|t\|)$$ • **Intervalo de Confianza**: $$IC \equiv \left[\widehat{\beta_J} \pm z_{\frac{\alpha}{2}}\, s(\widehat{\beta_J})\right]$$
Contraste de Hipótesis acerca de una	• **Hipótesis nula**: $$H_0: \beta_1 = \beta_2$$ $$H_0: \beta_1 - \beta_2 = 0$$

única combinación lineal de los parámetros	• **Hipótesis Alternativa**: $$H_1: \beta_1 < \beta_2$$ $$H_1: \beta_1 - \beta_2 < 0$$ • **Estadístico de Contraste**: $$t_{\widehat{\beta_j}} = \frac{(\widehat{\beta_1} - \widehat{\beta_2})}{s(\widehat{\beta_1} - \widehat{\beta_2})}$$ **Cálculo de $s(\widehat{\beta_1} - \widehat{\beta_2})$**: $$V(\widehat{\beta_1} - \widehat{\beta_2}) = V(\widehat{\beta_1}) + V(\widehat{\beta_2}) - 2C(\widehat{\beta_1}, \widehat{\beta_2})$$ $$s(\widehat{\beta_1} - \widehat{\beta_2}) = \left\{ [s(\widehat{\beta_1})]^2 + [s(\widehat{\beta_2})]^2 - 2s_{12} \right\}^{1/2}$$ • **Otra forma de cálculo**: $$Y = \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \beta_3 * X_3 + \varepsilon$$ $$H_0: \theta_1 = 0, siendo\ \theta_1 = \beta_1 - \beta_2$$ $$t_{\widehat{\beta_j}} = \frac{\theta_1}{s(\theta_1)}$$ Sustituir $\beta_1 = \theta_1 + \beta_2$ en la regresión original y estimamos: $$Y = \beta_0 + (\theta_1 + \beta_2) * X_1 + \beta_2 * X_2 + \beta_3 * X_3 + \varepsilon = \beta_0 + \theta_1 * X_1 + \beta_2(X_1 + X_2) + \beta_3 * X_3 + \varepsilon$$
Contraste de Restricciones Lineales Múltiples	• **Noción básica**: 1. El contraste de Restricciones Lineales Múltiples se utiliza para contrastar si un conjunto de variables independientes no ejerce un efecto parcial sobre la variable dependiente. 2. Modelo No Restringido: Aquel modelo inicial que contiene todas las variables explicativas. 3. Modelo Restringido: Aquel modelo resultado de haber sustituido las hipótesis nulas en el modelo inicial. • **Contraste de Restricciones de Exclusión**: **Modelo no restringido**: $$Y = \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \beta_3 * X_3 + \beta_4 * X_4 + \beta_5 * X_5 + \varepsilon$$ **Hipótesis nulas y alternativas**: $$H_0: \beta_3 = 0, \beta_4 = 0, \beta_5 = 0$$ $$H_1: H_0\ no\ es\ verdadera$$ **Modelo restingido**: $$Y = \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \varepsilon$$ **Estadístico de Contraste**: $$F = \frac{\frac{SCE_R - SCE_{NR}}{q}}{\frac{SCE_R}{n-k-1}} \rightarrow F_{q.n-k-1}$$

Como SCE_R no puede ser menor que SCE_{NR}, el estadístico F es siempre positivo.
En Econometría, todas las muestras son Asintóticas (n>30), y estos estadísticos convergen a una **distribución chi-cuadrado** χ^2: $q * F = w \rightarrow \chi^2_{\mathbf{q}}$

$$w = \frac{SCE_R - SCE_{NR}}{SCE_R} * n \rightarrow \chi^2_{\mathbf{q}}$$

Notación:
q= número de restricciones de exclusión de H_0.
k= número de variables explicativas del modelo no restringido.

Interpretación del Resultado de Contraste:

1. Si se rechaza H_0, decimos que $X_1, X_2, \dots, X_k$ son **estadísticamente significativa** de forma conjunta al nivel de significatividad adecuado.
2. Si no se rechaza H_0, entonces las variables son **conjuntamente no significativas**.

- **Relación entre los estadísticos t y F**:

Se puede demostrar que el estadístico F para contrastar la exclusión de una única variable es igual al **cuadrado** del estadístico t correspondiente.

$$F = t^2$$

O bien,

$$w = t^2$$

- **La forma R-cuadrado del estadístico F**:

1. Una razón para utilizar la forma de R-cuadrado se halla en que el R-cuadrado está siempre entre 0 y 1, mientras que las SCE pueden ser muy grandes dependiendo de las unidades de medida de Y.
2. Forma R-cuadrado del estadístico F:

$$F = \frac{\frac{R^2_{NR} - R^2_R}{q}}{\frac{1 - R^2_{NR}}{n-k-1}} \rightarrow F_{q.n-k-1}$$

En Econometría, todas las muestras son Asintóticas (n>30), y estos estadísticos convergen a una **distribución chi-cuadrado** χ^2: $q * F = w \rightarrow \chi^2_{\mathbf{q}}$

$$w = \frac{R^2_{NR} - R^2_R}{1 - R^2_{NR}} * n \rightarrow \chi^2_{\mathbf{q}}$$

3. Como $R^2_{NR} > R^2_R$, esto demuestra de nuevo que el estadístico F siempre será positivo.
4. Es importante no elevar al cuadrado el R-cuadrado antes de introducirlo en la fórmula, puesto que la elevación al cuadrado ya se ha realizado.

- **El Estadístico F para la significatividad conjunta de una regresión**:

Modelo no restringido:

$$Y = \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \beta_3 * X_3 + \beta_4 * X_4 + \beta_5 * X_5 + \varepsilon$$

Hipótesis nulas y alternativas:

$$H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$$
$$H_1: \beta_j \neq 0$$

Modelo restingido:

$$Y = \beta_0 + \varepsilon$$

Estadístico de Contraste:

$$F = \frac{\frac{R_{NR}^2}{k}}{\frac{1-R_{NR}^2}{n-k-1}} \rightarrow F_{q.n-k-1}$$

En este caso, R-cuadrado de $Y = \beta_0 + \varepsilon$ es igual a 0.

- **Contraste de Restricciones Lineales Generales**:

Modelo no restringido:

$$Y = \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \beta_3 * X_3 + \beta_4 * X_4 + \beta_5 * X_5 + \varepsilon$$

Hipótesis nulas y alternativas:

$$H_0: \beta_1 = 1;\ \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$$
$$H_1: \beta_j \neq 0$$

Modelo restingido:

$$Y = \beta_0 + X_1 + \varepsilon$$
$$Y - X_1 = \beta_0 + \varepsilon$$

Estadístico de Contraste:

$$F = \frac{\frac{SCE_R - SCE_{NR}}{q}}{\frac{SCE_R}{n-k-1}} \rightarrow F_{q.n-k-1}$$

En este caso, el modelo restringido es un modelo con un término constante, pero con una variable dependiente diferente de la del modelo no restingido. Por lo tanto, **no podemos utilizar la forma R-cuadrado del estadístico F**.

Como Regla General, deberá usarse la forma SCE del estadístico F si la regresión restringida presenta una variable dependiente distinta a la de la regresión no restringida.

Tema 4 Variables Binarias

Variables Binarias	
Introducción	• En la investigación econométrica suelen aparecer fenómenos de discriminación por razón de sexo, raza, nivel de estudios, lugar de residencia, etc, que no pueden analizarse utilizando únicamente variables cuantitativas. • Para analizar estos fenómenos discriminatorios, necesitaremos introducir información cualitativa en los modelos.
Concepto	• Las variables cualitativas se introducen en los modelos mediante variables binarias, ficticias o dummies. • Se trata de variables que toman únicamente dos valores: 1. Valor 1: Cuando el individuo o el elemento analizado cumpla una determinada característica. 2. Valor 0: Cuando el elemento o individuo analizado no cumple la citada característica. • Siempre se cumplirá que podamos definir tantas variables ficticias como niveles o categorías que tenga la variable cualitativa. • Siempre se cumplirá que la suma de las variables binarias de una variable cualitativa es igual a 1.
Método de Introducción de las Variables Binarias en el Modelo	• Si el modelo tiene término constante (β_0), tendremos que eliminar una variable binaria por cada variable cualitativa, porque de lo contrario aparecerá un problema de multicolinealidad perfecta, y el modelo no podrá estimarse. • Si el modelo no tiene término constante, podemos introducir todas las variables binarias. • El objetivo de las variables binarias es comparar regresiones de distintos colectivos y para ello, se pueden introduir las variables binarias de dos formas: 1. En forma aditiva: Cuando queramos analizar las posibles diferencias en los términos constantes de las regresiones de los distintos colectivos. 2. En forma de interacción: Cuando queramos analizar posibles diferencias en las pendientes de las regresiones de los distintos colectivos.
Modelos Básicos que utilizan Variables Binarias	
Modelo 1	• Para explicar la variable dependiente Y (Salario), utilizamos únicamente la variable cualitativa (Sexo) con variables binarias únicamente en forma aditiva. • En primer lugar, tendremos que definir las variables binarias correspondientes a las 2 categorías del sexo: $$X_1 = \begin{cases} 1 & mujer \\ 0 & hombre \end{cases} \;;\; X_2 = \begin{cases} 1 & hombre \\ 0 & mujer \end{cases}$$ • A continuación podremos plantear dos modelos alternativas: 1. **Modelo con Término Constante**: En este caso, deberíamos eliminar una variable binaria para evitar la multicolinealidad, por ejemplo, X_2: $$Y = \beta_0 + \beta_1 * X_1 + \varepsilon$$ En estos modelos, es especialmente importante la interpretación de coeficientes, así como los contrastes de discriminación, que sirven para verificar si existen diferencias entre los

distintos colectivos.

Tomando las esperanzas condicionadas de la ecuación anterior, vamos a obtener las ecuaciones salariables de mujeres y hombres:

- Salario Medio de Mujer:

$$E(Y/X_1 = 1) = \beta_0 + \beta_1 + E(\varepsilon/X_1 = 1) = \beta_0 + \beta_1$$

- Salario Medio de Hombre:

$$E(Y/X_1 = 0) = \beta_0 + \beta_1 * 0 + E(\varepsilon/X_1 = 0) = \beta_0$$

Restando las esperanzas condicionadas, obtenemos:

$$E(Y/X_1 = 1) - E(Y/X_1 = 0) = \beta_1$$

Interpretación de Coeficientes:

β_0: Salario Medio de los Hombres.

β_1: La diferencia salarial media entre mujeres y hombres.

Contraste:

Para contrastar si existe discriminación salarial entre mujeres y hombres:

$$H_0: \beta_1 = 0 \ (No\ hay\ discriminación)$$
$$H_1: \beta_1 \neq 0 \qquad (Hay\ discriminación)$$

2. **Modelo sin Término Constante**:

Al no incluir el término constante, podremos introducir las dos variables binarias, y la educación del modelo será:

$$Y = \delta_1 * X_1 + \delta_2 * X_2 + \varepsilon$$

Tomando esperanzas condicionadas en la ecuación anterior, obtenemos:

- Salario Medio Mujer:

$$E(Y/X_1 = 1, X_2 = 0) = \delta_1 * 1 + \delta_2 * 0 + E(\varepsilon/X_1 = 1,, X_2 = 0) = \delta_1$$

- Salario Medio Hombre:

$$E(Y/X_1 = 0, X_2 = 1) = \delta_1 * 0 + \delta_2 * 1 + E(\varepsilon/X_1 = 1,, X_2 = 0) = \delta_2$$

Interpretación de los Coeficientes:

δ_1 y δ_2 serán respectivamente los salarios medios de mujeres y hombres.

Contraste:

Para contrastar si existen diferencias salariales entre mujeres y hombres, habrá que plantear:

$$H_0: \delta_1 = \delta_2 \ (No\ hay\ discriminación)$$
$$H_1: \delta_1 \neq \delta_2 \qquad (Hay\ discriminación)$$

- Conclusión:
  1. Comparando el modelo con término constante y el modelo sin término constante, deducimos que en los modelos con término constante, el coeficiente de la variable binaria mide los **EFECTOS DIFERENCIALES**, mientras que en los modelos sin término constante, los coeficientes de las variables binarias miden **EFECTOS ABSOLUTOS**.

Modelo 2

- Para explicar la variable dependiente Y (Salario), utilizamos la variable cuantitativa años de estudio (X_1) y la variable cualitativa sexo, con variables binarias únicamente en forma aditiva.
- En primer lugar, definimos las dos variables binarias correspondientes al sexo:

$$X_2 = \begin{cases} 1 & mujer \\ 0 & hombre \end{cases} \ ; \ X_3 = \begin{cases} 1 & hombre \\ 0 & mujer \end{cases}$$

	• Vamos a centrarnos en el modelo con término constante y para plantearlo, eliminamos una de las variables binarias, por ejemplo, X_3. • La ecuación del modelo será la siguiente: $$Y = \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \varepsilon$$ • Tomando esperanzas condicionadas en la ecuación anterior, obtenemos las ecuaciones salariales de mujeres y hombres: 1. Salario Medio de Mujeres: $$E(Y/X_1, X_2 = 1) = \beta_0 + \beta_1 * X_1 + \beta_2 * 1 + E(\varepsilon/X_1, X_2 = 1) = \beta_0 + \beta_1 * X_1 + \beta_2 = (\beta_0 + \beta_2) + \beta_1 * X_1$$ 2. Salario Medio de Hombres: $$E(Y/X_1, X_2 = 0) = \beta_0 + \beta_1 * X_1 + \beta_2 * 0 + E(\varepsilon/X_1, X_2 = 0) = \beta_0 + \beta_1 * X_1$$ 3. Restando las esperanzas condicionadas: $$E(Y/X_1, X_2 = 1) - E(Y/X_1, X_2 = 0) = \beta_2$$ • Interpretación de los Coeficientes: 1. β_0: Salario medio de los hombres sin estudios ($X_1 = 0$). 2. β_1: Pendiente de los dos colectivos, y decimos que a un incremento de un año en los estudios, le corresponde, en promedio, una variación de β_1 unidades monetarias en los salarios de mujeres y hombres. 3. β_2: Diferencia salarial media entre mujeres y hombres a igualdad de años de estudios, o manteniendo X_1 en ceteris paribus. • Contraste: Para verificar si existen diferencias salariales entre mujeres y hombres: $$H_0: \beta_2 = 0\ (No\ hay\ diferencias)$$ $$H_1: \beta_2 \neq 0\ \ (Hay\ diferencias)$$
Modelo 3	• Para explicar la variable dependiente Y (Salario), utilizamos la variable cuantitativa años de estudio (X_1) y la variable cualitativa sexo, con variables binarias únicamente en forma aditiva. • En primer lugar, definimos las dos variables binarias correspondientes al sexo: $$X_2 = \begin{cases} 1 & mujer \\ 0 & hombre \end{cases} \ ;\ X_3 = \begin{cases} 1 & hombre \\ 0 & mujer \end{cases}$$ • Vamos a centrarnos en el modelo con término constante y para plantearlo, eliminamos una de las variables binarias, por ejemplo, X_3. • La ecuación del modelo será la siguiente: $$Y = \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \beta_3 * X_1 * X_2 + \varepsilon$$ • Al incluir la variable binaria X_2 en forma aditiva y de interacción con X_1, este modelo permite analizar diferencias en el término constante y la pendiente de las regresiones de mujeres y hombres. • Tomando esperanzas condicionadas en la ecuación anterior, obtenemos las ecuaciones salariales de mujeres y hombres. 1. Salario Medio de Mujeres: $$E(Y/X_1, X_2 = 1) = \beta_0 + \beta_1 * X_1 + \beta_2 * 1 + \beta_3 * X_1 * 1 + E(\varepsilon/X_1, X_2 = 1) = \beta_0 + \beta_1 * X_1 + \beta_2 + \beta_3 * X_1 = (\beta_0 + \beta_2) + (\beta_1 + \beta_3) * X_1$$

	2. Salario Medio de Hombres: $E(Y/X_1, X_2 = 0) = \beta_0 + \beta_1 * X_1 + \beta_2 * 0 + \beta_3 * X_1 * 0 + E(\varepsilon/X_1, X_2 = 0)$ $= \beta_0 + \beta_1 * X_1$ 3. Restando las esperanzas condicionadas: $E(Y/X_1, X_2 = 1) - E(Y/X_1, X_2 = 0) = \beta_2 + \beta_3 * X_1$ • Interpretación de los Coeficientes: 1. β_0: Salario medio de los hombres sin estudios ($X_1 = 0$). 2. β_1: A un incremento de un año en los estudios, le corresponde, en promedio, una variación de β_1 unidades monetarias en el salario de los hombres. 3. β_2: La diferencia salarial media entre mujeres y hombres a igualdad de años de estudio. 4. â$_3$: La diferencia salarial media entre mujeres y hombres cuando ambos incrementan un año en sus estudios. • Además, en los modelos con términos de interacción que incluyen variables binarias, pueden pedir el EFECTO PARCIAL de la variable explicativa que acompaña a la variable binaria. Por ejemplo, para hallar el efecto parcial del rendimiento de la educación (X_1): $\frac{\Delta Y}{\Delta X_1} = \beta_1 + \beta_3 * X_2$ Lógicamente, este efecto no es constante. • Podríamos contrastar si el efecto parcial de la educación depende del sexo, planteando lo siguiente: $H_0: \beta_3 = 0$ $H_1: \beta_3 \neq 0$ • También se puede plantear otros contrastes como los siguientes: 1. Para contrastar si existen diferencias en el término constante de las regresiones de mujeres y hombres, planteamos: $H_0: \beta_2 = 0$ $H_1: \beta_2 \neq 0$ 2. Para contrastar si existen diferencias en la pendiente de las regresiones de mujeres y hombres: $H_0: \beta_3 = 0$ $H_1: \beta_3 \neq 0$ 3. Para contrastar si existen diferencias salariales entre mujeres y hombres: $H_0: \beta_2 = \beta_3 = 0$ $H_1: \beta_2 \neq 0 \ y/o \ \beta_3 \neq 0$
Observaciones	• Además de estos 3 modelos básicos, se pueden plantear modelos más complejos que incluyan 2 o más variables cualitativas e incluso podrán aparecer términos de interacción entre 2 variables binarias. • En cualquier caso, siempre tendremos en cuenta que si el modelo tiene término constante, tendremos que eliminar una variable binaria por cada variable cualitativa para evitar la multicolinealidad. • Una vez planteado el modelo, calcularemos las diferencias entre colectivos tomando esperanzas condicionadas en el modelo poblacional.

<table>
<tr><th colspan="2">Modelo de Diferencias en Diferencias</th></tr>
<tr><td>Modelo de Diferencias en Diferencias</td><td>

- Hasta ahora, todos los modelos planteados se estimaban apartir de datos de sección cruzada, que eran conjuntos de observaciones de una o más variables medidas en distintas unidades económicas en un instante de tiempo. Por ejemplo, la renta de un conjunto de individuos en 2009.
- Se denominan datos de sección cruzada fusionadas o datos de panel, a conjuntos de observaciones de una o más variables medidas en distintas unidades económicas en dos períodos distintos de tiempo entre los que ha habido algún cambio estructural político, económico o social. Por ejemplo, el precio de un conjunto de viviendas en 2005 y en 2010 entre los que ha habido un cambio estructural, una crisis económica.
- Para analizar los efectos de este cambio estructural, se plantea el modelo de diferencias en diferencias.

</td></tr>
<tr><td>Aplicación</td><td>

- Nuestro objetivo es evaluar la repercusión de una elevación del tope máximo en el subsidio de desempleo sobre la duración.
- Definición de las Variables:
  Y : Duración del desempleo en semanas (medida en logaritmos).
  $X_1 = \begin{cases} 1 & \textit{después del cambio de política} \\ 0 & \textit{antes del cambio de la política} \end{cases}$
  $X_2 = \begin{cases} 1 & \textit{renta alta} \\ 0 & \textit{renta baja} \end{cases}$
- Escribimos el modelo en términos poblacionales:
  $$Y = \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \beta_3 * X_1 * X_2 + \varepsilon$$
- La ecuación anterior es un modelo de diferencias en diferencias, y tiene como objetivo, analizar cómo y a quién afectará el cambio de polícia económica, que en este caso, es la elevación del tope máximo del subsidio por baja laboral. A priori, esta medida sólo debería afectar a los individuos de rentas altas, que perciben el mayor subsidio y no debería afectar a los de rentas bajas.
- Comenzamos interpretando los coeficientes de la ecuación anterior:
  1. β_0: La duración de la baja laboral en logatirmos para los individuos de renta no alta antes del cambio.
  2. β_1: La diferencia en la duración de la baja entre antes y después del cambio para los individuos de rentas no altas.
  3. β_2: La diferencia en la duración de la baja laboral entre los individuos de renta alta y no alta antes del cambio.
  4. β_3: Mide el efecto causal del cambio de política, es decir, la diferencia en la duración de la baja laboral de los individuos de rentas altas después del cambio.

</td></tr>
</table>

Tema 5 Errores de Especificación

Omisión de Variables Relevantes (Subespecificación del Modelo)

Concepto

- Consisite en eliminar del modelo alguna variable explicativa significativa.
- La exclusión de una variable relevante o de subespecificación del modelo.

Omisión de Variables Relevantes

- Consideremos el siguiente modelo correctamente especificado, que cumple todos los supuestos:

$$Y = \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \varepsilon$$

- Sin embargo, por falta de información, decidimos omitir la variable X_2, que es relevante (significativa) porque $\beta_2 \neq 0$. Por eso, planteamos el siguiente modelo mal especificado:

$$Y = \delta_0 + \delta_1 * X_1 + \varepsilon$$

- Vamos a demostrar que la omisión de la variable relevante X_2 hace que en el modelo segundo la $\boldsymbol{E(\varepsilon/X_1) \neq 0}$.
- Teniendo en cuenta que el error:

$$U = \beta_2 * X_2 + \varepsilon$$

- Tomando esperanzas condicionadas:

$$E(U/X_1) = E(\beta_2 * X_2 + \varepsilon/X_1) = \beta_2 E(X_2/X_1) + E(\varepsilon/X_1) = \beta_2 E(X_2/X_1) \neq 0$$

- El incumplimiento del supuesto $E(U/X_1) = 0$ provoca que los estimadores de MCO del modelo segundo sean **SESGADOS** e **INCONSISTENTES**.

Sesgo por Omisión de Variable

- Para obtener el sesgo, necesitamos una expresión que relacione los parámetros de los dos modelos anteriores.
- Para ello, partimos de la expresión de δ_1 y en ella, sustituimos la variable Y por el verdadero modelo:

$$\boldsymbol{\delta_1} = \frac{C(X_1, Y)}{V(X_1)} = \frac{C(X_1, \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \varepsilon)}{V(X_1)}$$

$$= \frac{C(X_1, \beta_0) + \beta_1 * C(X_1, X_1) + \beta_2 * C(X_1, X_2) + C(X_1, \varepsilon)}{V(X_1)}$$

$$= \frac{\beta_1 * V(X_1) + \beta_2 * C(X_1, X_2)}{V(X_1)} = \boldsymbol{\beta_1 + \beta_2 * \frac{C(X_1, X_2)}{V(X_1)}}$$

- $\boldsymbol{\delta_1}$ no es igual a $\boldsymbol{\beta_1}$ por regla general: $\boldsymbol{\delta_1}$ **es un estimador sesgado de** $\boldsymbol{\beta_1}$.
- El cociente que multiplica $\boldsymbol{\beta_2}$ tiene una interpretación simple: Es simplemente el coeficiente de la pendiente de la regresión de X_2 sobre X_1.

$$X_2 = \gamma_0 + \gamma_1 * X_1 \text{, } siendo\ \gamma_1 = \frac{C(X_1, X_2)}{V(X_1)}$$

- Por lo tanto, podemos expresar de la siguiente manera:

$$\boldsymbol{\delta_1 = \beta_1 + \beta_2 * \gamma_1}$$

- El sesgo en $\boldsymbol{\delta_1}$ es $\boldsymbol{\delta_1 - \beta_1 = \beta_2 * \gamma_1}$. Esto es lo que normalmente se denomina **sesgo por omisión de variable**.
- Resumen del sesgo en $\boldsymbol{\beta_1}$ cuando se omite X_2:

	$\boldsymbol{C(X_1, X_2)} > 0$	$\boldsymbol{C(X_1, X_2)} < 0$
$\boldsymbol{\beta_2} > 0$	Sesgo Positivo	Sesgo Negativo

<table>
<tr><td></td><td>
<table>
<tr><td></td><td>(Sobreestimación)</td><td>(Infraestimación)</td></tr>
<tr><td>$\boldsymbol{\beta_2} < 0$</td><td>Sesgo Negativo
(Infraestimación)</td><td>Sesgo Positivo
(Sobreestimación)</td></tr>
</table>

- Si $\boldsymbol{\delta_1} > \boldsymbol{\beta_1}$, decimos que $\boldsymbol{\delta_1}$ tiene un **sesgo al alza**. Si $\boldsymbol{\delta_1} < \boldsymbol{\beta_1}$, decimos que $\boldsymbol{\delta_1}$ tiene un **sesgo a la baja**.
- De forma general, se puede calcular el sesgo mediante la siguiente fórmula:

$$\boldsymbol{Sesgo = Coef.Variable\ Omitida} * \frac{\boldsymbol{C(Variable\ incluida, Variable\ Omitida)}}{\boldsymbol{V(Variable\ incluida)}}$$
</td></tr>
<tr><td>Conclusión</td><td>

- La omisión de una variable relevante provoca SESGO e INCONSISTENCIA en el estimador de MCO excepto en el caso de incorrelación entre las variables omitidas y las variables incluidas.
- V(δ_1) siempre es más pequeña que V(β_1).
</td></tr>
<tr><td colspan="2">Inclusión de Variables Irrelevantes (Sobreespecificación del Modelo)</td></tr>
<tr><td>Concepto</td><td>

- Significa que una o más de las variables independientes que se han incluido en el modelo no tiene ningún efecto parcial sobre Y en la población, es decir, que su coeficiente poblacional es nulo.
</td></tr>
<tr><td>Inclusión de Variables Irrelevantes</td><td>

- Supongamos que establecemos un modelo del tipo:

$$Y = \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \beta_3 * X_3 + \varepsilon$$

- Sin embargo, X_3 no tiene ningún efecto sobre Y cuando X_1 y X_2 han sido tenidos en cuenta, lo que significa que $\beta_3 = 0$. En términos de Esperanzas Condicionadas, $\boldsymbol{E(Y/X_1, X_2, X_3)} = \boldsymbol{E(Y/X_1, X_2)} = \beta_0 + \beta_1 * X_1 + \beta_2 * X_2$.
</td></tr>
<tr><td>Efectos</td><td>

- En cuanto a la insesgadez de β_1 y β_2, no tiene ningún efecto.
- El añadir una variable irrelevante a una ecuación normalmente produce un aumento de las varianzas del resto de los estimadores MCO debido a la multicolinealidad.
</td></tr>
<tr><td>Conclusión</td><td>

- El incluir una o más variables irrelevantes en un modelo de regresión múltiple, o el sobreespecificar el modelo, no afecta a la insesgadez de los estimadores MCO.
- Al producir un incremento en las varianzas de dichos estimadores, lo que se conoce como Pérdida de Eficiencia, como consecuencia de ello, la inferencia seguirá siendo válida, pero escasamente fiable.
- Pese a ello, la inclusión de variables irrelevantes tiene efectos menos graves que la omisión de variables relevantes.
</td></tr>
<tr><td colspan="2">Errores de Medida en las Variables</td></tr>
<tr><td>Concepto</td><td>

- Cuando planteemos un modelo, puede ocurrir que la variable dependiente o alguna variable independiente estén medidas con error.
- Especialmente grave es el error de medida en las variables explicativas, ya que provoca el SESGO y la INCONSISTENCIA de los estimadores de MCO.
</td></tr>
<tr><td>Error de Medida en la Variable Dependiente</td><td>

- Supongamos el siguiente modelo:

$$Y^* = \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \varepsilon$$

- Este modelo no puede estimar, porque la variable dependiente Y^* no es observable, y en su lugar, observamos la variable Y, que cumple la siguiente ecuación:
</td></tr>
</table>

	$$\boldsymbol{Y = Y^* + v}\,, \textit{siendo } \boldsymbol{v} \textit{ el error de medida.}$$ • Nuestro objetivo es poder estimar el modelo, y para ello, despejamos Y^* en la ecuación anterior y lo sustituimos en la ecuación primera: $$\boldsymbol{Y^* = Y - v}$$ $$Y - v = \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \varepsilon$$ $$Y = \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \varepsilon + v$$ • Llamamos a $\varepsilon + v = u$, luego: $$Y = \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + u$$ • A continuación, vamos a comprobar los efectos del error de medición en Y y para ello, realizamos dos supuestos adicionales: 1. $Cov(\varepsilon, v) = 0$ 2. Supuesto clásico del Error de Medida en la Variable Dependiente: Suponemos que las variables explicativas X_1 y X_2 están incorrelacionadas con el error de medida v. Es decir: $Cov(X_1, v) = 0$ y $Cov(X_2, v) = 0$. • Efectos de Error de Medida en la Variable Dependiente: 1. Se puede demostrar que si estimamos $Y = \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + u$ por MCO, obtendremos estimadores INSESGADOS y CONSISTENTES. 2. Aparentemente, el error de medida en la variable dependiente no tendrá consecuencias graves, pero no es así, porque el error del modelo ha aumentado, ha pasado de ε a u. Además, la varianza de dicho error ha aumentado: $$V(u) = V(\varepsilon) + V(v) + 2C(\varepsilon, v) = V(\varepsilon) + V(v) > V(\varepsilon)$$ Este incremento en la varianza del error provoca un incremento en las varianzas de los estimadores, lo que supone una Pérdida de Eficiencia, y como consecuencia, la inferencia sigue siendo válida pero poco fiable. 3. Estos efectos son idénticos a los de la inclusión de variables irrelevantes. • Caso especial en el error de medida en la variable dependiente: 1. Supongamos que $Cov(X_1, v) \neq 0$, entonces se incumple el supuesto clásico de error de medida en la variable dependiente. 2. Lógicamente, v va a formar parte de u, por lo que $Cov(X_1, u) \neq 0$. 3. Por lo tanto, la consecuencia será que $E(u/X_1, X_2) \neq 0$.
Error de Medida en la Variable Independiente	• Supongamos en el siguiente modelo: $$Y = \beta_0 + \beta_1 * X^* + \varepsilon$$ • Como X^* no es observable, el modelo no se puede estimar, y en su lugar, observamos X que cumple la siguiente ecuación: $$\boldsymbol{X = X^* + v}, \textit{siendo } \boldsymbol{v} \textit{ el error de medida}$$ • Además, antes de analizar los efectos de este error de medida, efectuamos los siguientes supuestos: 1. $\boldsymbol{Cov(X, \varepsilon) = 0}$ 2. Supuesto Clásico del error de medida en la variable independiente: Suponemos que la variable medida con error está incorrelacionada con el error de medida v: $$\boldsymbol{Cov(X^*, v) = 0}$$

- Para poder estimar la ecuación, en primer lugar despejamos X^* en la ecuación segunda, y lo sustituimos en la ecuación primera:

$$X^* = X - v$$
$$Y = \beta_0 + \beta_1 * (X - v) + \varepsilon$$
$$Y = \beta_0 + \beta_1 * X - \beta_1 * v + \varepsilon$$
$$Y = \beta_0 + \beta_1 * X + \varepsilon - \beta_1 * v$$

- Llamamos a $u = \varepsilon - \beta_1 * v$, luego el modelo con todas las variables observables será:

$$Y = \beta_0 + \beta_1 * X + u$$

- Si estimamor la ecuación anterior por MCO, se produce un incremento en la varianza del error que provoca que aumenten también las varianzas de los estimadores.
- Sin embargo, el efecto realmente importante, es el SESGO e INCONSISTENCIA de los estimadores de MCO.
- Para demostrar este sesgo, partimos de la expresión de $\widehat{\beta_1}$:

$$\widehat{\beta_1} = \frac{C(X,Y)}{V(X)} = \frac{C(X, \beta_0 + \beta_1 * X + u)}{V(X)} = \frac{C(X,\beta_0) + \beta_1 * C(X,X) + C(X,u)}{V(X)}$$
$$= \frac{\beta_1 * V(X) + C(X,u)}{V(X)} = \beta_1 + \frac{C(X,u)}{V(X)}$$
$$C(X,u) = C(X, \varepsilon - \beta_1 * v) = C(X,\varepsilon) - \beta_1 C(X,v) = -\beta_1 C(X,v) = -\beta_1 * C(X^* + v, v)$$
$$= -\beta_1 * [C(X^*, v) + C(v,v)] = -\beta_1 * V(v)$$
$$V(X) = V(X^* + v) = V(X^*) + V(v) + 2C(X^*, v) = V(X^*) + V(v)$$
$$Sesgo\ de\ \widehat{\beta_1} = \frac{C(X,u)}{V(X)} = \frac{-\beta_1 * V(v)}{V(X^*) + V(v)}$$

- A continuación, vamos a demostrar que el error de medida en X provoca que $\widehat{\beta_1}$ sea un estimador inconsistente de β_1, es decir, $p\lim_{n\to\infty} \widehat{\beta_1} \neq \beta_1$.

$$p\lim_{n\to\infty}\left(\widehat{\beta_1}\right) = p\lim_{n\to\infty}\left(\beta_1 + \frac{C(X,u)}{V(X)}\right) = p\lim_{n\to\infty}(\beta_1) + \frac{p\lim_{n\to\infty} C(X,u)}{p\lim_{n\to\infty} V(X)} = \beta_1 + \frac{C(X,u)}{V(X)} \neq \beta_1$$

Sesgo Asintótico o Sesgo de Inconsistencia:

$$Sesgo\ Asintótico\left(\widehat{\beta_1}\right) = p\lim_{n\to\infty} \widehat{\beta_1} - \beta_1$$
$$Sesgo\ Asintótico\left(\widehat{\beta_1}\right) = \frac{C(X,u)}{V(X)}$$

- Conclusiones sobre el sesgo:
  1. A la vista de la expresión obtenida, deducimos que si β_1 es positivo, el sesgo será negativo.
  2. La magnitud del sesgo depende directamente de la varianza del error de medida, $V(v)$, e inversamente de la varianza de la variable medida con errores $V(X^*)$.
  3. Existen dos casos especiales en los que el sesgo es nulo y $\widehat{\beta_1}$ es consistente:
     a) Si X tiene un error de medida nulo ($v = 0$), entonces $X = X^*$ y el sesgo será 0.
     b) Supongamos que todas las observaciones de X estuviesen medidas con el

	mismo error, es decir $v_1 = v_2 = \cdots = v_k$, en ese caso, $V(v) = 0$, y el sesgo también será nulo. • Conclusión: 1. El error de medida en las variables explicativas provoca sesgo e inconsistencia en todos los estimadores de MCO, excepto en los casos especiales. 2. La causa del sesgo es que la variable medida con error (X^*) está correlacionada con el error del modelo inicial (ε), es decir, C(X^*,ε) $\neq 0$. 3. En términos precisos, se dice que X^* es una variable endógena. 4. El sesgo y la inconsistencia de los estimadores de MCO hace que la inferencia esté invalidada.

Tema 6 Modelos con Variables Explicativas Endógenas

Estimación de Variables Instrumentales	
Estimación de Variables Instrumentales	• Antes de plantear este método de estimación debemos definir una variable instrumental e instrumento Z, que es una variable que cumple las dos siguientes condiciones: 1. La variable instrumental Z debe estar incorrelada con el término de error del modelo ε, es decir, $\boldsymbol{C(Z,\varepsilon)=0}$. Esta condición no puede contrastarse, y sólo puede razonarse de manera terórica. 2. La variable instrumental Z debe estar correlacionada con la variable endógena X. Es decir, $\boldsymbol{C(Z,X)\neq 0}$. Esta segunda condición de instrumento puede comprobarse planteando una regresión de la variable endógena X frente a la variable instrumental Z. $$X=\pi_0+\pi_1 Z+v$$ Planteamos el contraste de identificación: $$H_0:\pi_1=0$$ $$H_1:\pi_1\neq 0$$ Si $\pi_1\neq 0$, por lo que X y Z estarán correlacionadas y se cumplirá la segnda condición de instrumento. • Una vez comprobadas las dos condiciones de instrumento, si se cumplen, entonces será un instrumento válido o adecuado. • Para ello, introducimos en la C(Z,Y) el valor de la variable Y de la ecuación: $$C(Z,Y)=C(Z,\beta_0+\beta_1*X+\varepsilon)=C(Z,\beta_0)+\beta_1 C(Z,X)+C(Z,\varepsilon)=\beta_1 C(Z,X)$$ $$\beta_1=\frac{C(Z,Y)}{C(Z,X)}$$ • Aplicando el principio de analogía, obtenemos la expresión del estimador de variables instrumentales: $$\widehat{\beta_1}_{(VI)}=\frac{\hat{C}(Z,Y)}{\hat{C}(Z,X)}=\frac{S_{Z,Y}}{S_{Z,X}}$$ • Se puede demostrar que $\widehat{\beta_1}_{(VI)}$ es un estimador INSESGADO y CONSISTENTE de β_1. Insesgadez: $E\left(\widehat{\beta_1}_{(VI)}\right)=\beta_1$ Consistencia: $p\lim_{n\to\infty}\left(\widehat{\beta_1}_{(VI)}\right)=\beta_1$ • Sin embargo, a pesar de ser un estimador insesgado y consistente, ocurre que: $$V\left(\widehat{\beta_1}_{(MCO)}\right)=\frac{\delta^2}{n*s_X^2}<V\left(\widehat{\beta_1}_{(VI)}\right)=\frac{\delta^2}{n*s_X^2*R_{X,Z}^2}$$ Donde $R_{X,Z}^2$ es el coeficiente de determinación de la regresión de X sobre Z. • Conclusión: 1. El estimador de MCO es siempre más eficiente (tiene menor varianza) que el estimador de variable instrumental. 2. Cuanto mayor sea la relación entre X y Z, menor varianza tendrá el estimador de

	variables instrumentales. 3. Bajo endogeneidad, utilizaremos el estimador de variables instrumentales.
Estimación por Mínimos Cuadrados Bietápicos o en dos etapas (MC2E)	
Estimación por Mínimos Cuadrados Bietápicos o en dos etapas	• Hasta ahora, hemos analizado el problema de la endogeneidad únicamente en el Modelo de Regresión Simple. La solución de la estimación de Variables Instrumentales, que producía estimadores de Variables Instrumentales, que producía estimadores insesgados y consistentes. • Sin embargo, el método de Variables Instrumentales, es complejo de aplicar en el modelo de regresión múltiple, y en su lugar utilizamos la estimación de mínimos cuadrados bietápicos. • Supongamos el siguiente modelo: $$Y_1 = \beta_0 + \beta_1 * Y_2 + \beta_2 * Z_1 + u$$ Donde Y_2 es una **variable endógena**, es decir, $C(Y_2, u) \neq 0$ Z_1 es una **variable exógena**, es decir, $C(Z_1, u) = 0$ • Si estimamos esta ecuación por MCO, obtendríamos estimadores sesgados e inconsistentes, debido a la endogeneidad de Y_2. Para obtener estimadores insesgados y consistentes de β_0, β_1 y β_2, utilizamos la estimación de Mínimos cuadrados bietápicos o en dos etapas.
Caso 1	• **Disponemos de una sola variable endógena externa $\boldsymbol{Z_2}$.** • Buscamos el mejor instrumento posible para la variable endógena Y_2, y dicho instrumento es una combinación lineal de todas las variables exógenas internas y externas al modelo. $$Y_2 = \pi_0 + \pi_1 * Z_1 + \pi_2 * Z_2 + v$$ • **1ªEtapa**: Estimamos por MCO la ecuación: $$\widehat{Y_2} = \hat{\pi}_0 + \hat{\pi}_1 * Z_1 + \hat{\pi}_2 * Z_2$$ El objetivo de esta etapa es determinar si existe relación o no entre la variale endógena externa o posible instrumento. Entonces planteamos el contraste de identificación: $$H_0: \pi_2 = 0$$ $$H_1: \pi_2 \neq 0$$ En este contraste, debemos rechazar H_0 y decimos que se cumplirá la condición de identificación ($\pi_2 \neq 0$). Entonces π_2 será un instrumento válido. • **2ªEtapa**: En la ecuación $Y_1 = \beta_0 + \beta_1 * Y_2 + \beta_2 * Z_1 + u$, sustituimos la variable endógena Y_2, por los valores ajustados $\widehat{Y_2}$. $$Y_1 = \beta_0 + \beta_1 * \widehat{Y_2} + \beta_2 * Z_1 + u$$ Si estimamos la ecuación anterior por MCO, obtenemos los estimadores de mínimos cuadrados bietápicos, $\widehat{\beta_0}_{(MC2E)}, \widehat{\beta_1}_{(MC2E)}, \widehat{\beta_2}_{(MC2E)}$. Se trata de estimadores consistentes y asintóticamente normales, es decir, en muestras asintóticas, tienen distribución normal. Sin embargo, al igual que ocurría en Variables Instrumentales, los estimadores de Mínimos cuadrados bietápicos, son menos eficientes (tienen mayor varianza que los de Mínimos Cuadrados Ordinarios).

Caso 2	• **Disponemos de más de 1 instrumento**, es decir, de más de 1 variable exógena externa. Suponemos que tenía 2 variables exógenas externas: Z_2 y Z_3. • Buscamos el mejor instrumento posible para Y_2 que es una combinación lineal de todas las variables exógenas internas (Z_1) y externas (Z_2 y Z_3). $$Y_2 = \pi_0 + \pi_1 * Z_1 + \pi_2 * Z_2 + \pi_3 * Z_3 + v$$ • **1ªEtapa**: Estimamos por MCO la ecuación: $$\widehat{Y_2} = \hat{\pi}_0 + \hat{\pi}_1 * Z_1 + \hat{\pi}_2 * Z_2 + \hat{\pi}_3 * Z_3$$ A continuación planteamos el contraste de identificación que sirve para determinar si los instrumentos Z_2 y Z_3 están o no correlacionados con la variable endógena Y_2. $$H_0: \pi_2 = \pi_3 = 0$$ $$H_1: \pi_2 \neq 0; \pi_3 \neq 0$$ Si $\pi_2 \neq 0; \pi_3 \neq 0$, se cumplirá la condición de identificación, y los instrumentos Z_2 y Z_3 están correlacionados con Y_2. Por lo tanto, son instrumentos válidos. • **2ªEtapa**: En En la ecuación $Y_1 = \beta_0 + \beta_1 * Y_2 + \beta_2 * Z_1 + u$, sustituimos la variable endógena Y_2, por los valores ajustados $\widehat{Y_2}$. $$Y_1 = \beta_0 + \beta_1 * \widehat{Y_2} + \beta_2 * Z_1 + u$$ Si estimamos la ecuación anterior por MCO, obtenemos los estimadores de Mínimos Cuadrados Bietápicos o en 2 etapas, que serán **consistentes** y **asintóticamente normales**, siempre que los instrumentos Z_2 y Z_3 sean válidos.
Conclusión	• En general, **si los instrumentos no son válidos** por incumplir alguna de las condiciones de instrumento, los estimadores de Mínimos Cuadrados Bietápicos o en 2 etapas serán **INCONSISTENTES** y **peores que los de MCO**, puesto que los estimadores de MCO también son inconsistentes, sin embargo, tienen mínima varianza.
Contraste de Sobreidentificación de SARGAN	
Contraste de Sobreidentificación de SARGAN	• Planteamos una regresión de los residuos de Mínimos Cuadrados Bietápicos o en 2 etapas, llamados $\tilde{u}$ frente a todas las variables exógenas internas y externas. $$\tilde{u} = \delta_0 + \delta_1 Z_1 + \delta_2 Z_2 + \delta_3 Z_3 + \delta_4 Z_4 + u$$ • Las hipótesis del contraste de Sargan son: $$H_0: todos\ los\ instrumentos\ están\ incorrelados\ con\ el\ error$$ $$H_1: algún\ instrumento\ está\ correlacionadas\ con\ el\ error$$ El estadístico de Contraste es: $$n * R^2_{\tilde{u}} \sim \chi_{q-r}$$ • Si todos los instrumentos están incorrelacionados con el error, entonces serán instrumentos válidos. • Si algún instrumento está correlacionado con el error, entonces no será válido. • Sin embargo, el contraste de Sargan no dice qué instrumento es no válido.