

# Econometría Aplicada

## Propensity Score Matching

### Aplicación

Edinson Tolentino

MSc Economics

email: [edinson.tolentino@gmail.com](mailto:edinson.tolentino@gmail.com)

Twitter: [@edutoleraymondi](https://twitter.com/edutoleraymondi)

Educate Peru

# Contenido



Introducción

Pregunta 1

Pregunta 2

Pregunta 3

Pregunta 4

Pregunta 5

# Introducción



- Analisis de la productividad laboral y el crimen

## La Información

La base de datos es trabajada en base a la Encuesta Nacional de Empresas (ENE-2015)

Se busca analizar, el rol del cual el crimen afecta a la productividad

### Cuadro: Descripción de variables

Variables	Descripción
<i>crimen1</i>	== 1 , si la empresa sufrio crimen (dinero), = 0 otro caso
<i>crimen2</i>	== 1 , si la empresa sufrio crimen (activos), = 0 otro caso
<i>female</i>	== 1 , si la empresas es conducida por mujer, = 0 otro caso
<i>Edad</i>	edad de la empresa
<i>SF</i>	== 1, si la empresa acceso de credito
<i>edadgroup1</i>	== 1, si el conductor es menor 30 años, = 0 otro caso
<i>edadgroup2</i>	== 1, si el conductor de 30-36 años, = 0 otro caso
<i>edadgroup3</i>	== 1, si el conductor de 36-46 años, = 0 otro caso
<i>edadgroup4</i>	== 1, si el conductor es mayor 55 años, = 0 otro caso
<i>exporta</i>	== 1, si la empresa exporta, = 0 otro caso
<i>solefirms</i>	== 1, si la empresa es propia
<i>own house</i>	== 1, si el agricultor posee residencia propia
<i>manufacturing</i>	== 1, si la empresa pertenece manufactura
<i>commerce</i>	== 1, si la empresa pertenece comercio
<i>services</i>	== 1, si la empresa pertenece servicios

## Pregunta 1



## Pregunta 1

- **Pregunta 1: Examine la información separada si la empresa sufre de un crimen o no. ¿Qué podemos concluir?**

## Pregunta 1

- ▶ **Pregunta 1: Examine la información separada si la empresa sufre de un crimen o no. ¿Qué podemos concluir?**
- ▶ Se realiza un analisis descriptivo de las variables demograficas del conductor, características de empresas (acceso al credito, acceso a exportar, sector economico , etc)

## Pregunta 1

- ▶ **Pregunta 1: Examine la información separada si la empresa sufre de un crimen o no. ¿Qué podemos concluir?**
- ▶ Se realiza un análisis descriptivo de las variables demográficas del conductor, características de empresas (acceso al crédito, acceso a exportar, sector económico, etc)
- ▶ En promedio, muchas de las variables no son estadísticamente significativas (evidenciando) el rol de ser variables que observables para el matching.



# Pregunta 1

**Cuadro:** Estadística descriptiva (promedio)

	Tratado (Crimen)	Controles (No-crimen)	Diferencia
gender condition	0.19 (0.39)	0.22 (0.42)	0.03
30-36 years old (owned)	0.09 (0.29)	0.08 (0.28)	-0.01
36-46 years old (owned)	0.30 (0.46)	0.31 (0.46)	0.00
44-56 years old (owned)	0.36 (0.48)	0.37 (0.48)	0.01
More 55 years old (owned)	0.24 (0.43)	0.26 (0.44)	0.02
Access to capital credit	0.48 (0.50)	0.40 (0.49)	-0.08**
Years firms	12.43 (6.15)	12.27 (6.04)	-0.16
export status	0.11 (0.31)	0.10 (0.30)	-0.01
Sole firms (propietario)	0.38 (0.49)	0.40 (0.49)	0.02
sector2r4_2==Industrias manufactureras	0.26 (0.44)	0.34 (0.48)	0.09***
sector2r4_2==Comercio	0.10 (0.29)	0.09 (0.28)	-0.01
sector2r4_2==Servicios	0.37 (0.48)	0.36 (0.48)	-0.01
Observations	407	5708	6115

## Pregunta 1

Modelo de regresion (MCO) del efecto crimen sobre el log-productividad ( $\log(\text{productividad})$ )

**Cuadro:** Modelo de Regresion

	Model (1)	Model (2)
==1 crime (Capital)	-0.01 (0.06)	-0.08 (0.05)
Observaciones	6115	6115
Adj. R <sup>2</sup>	-0.00016	0.14
Sectors FE		✓
Controls		✓

Fuente: INEI - ENE.

Elaboracion: Autor

\*\*\*, \*\*, \* denote statistical significance at the 1 %, 5 % and 10 % levels respectively for zero.

## Pregunta 2



## Pregunta 2



- ▶ Ahora estimamos el modelo de regresión logístico:

## Pregunta 2

- ▶ Ahora estimamos el modelo de regresión logístico:

$$\text{prob}(\text{crimen} == 1) = \frac{\exp(z_i)}{1 + \exp(z_i)}$$

## Pregunta 2

- Ahora estimamos el modelo de regresión logístico:

$$\text{prob}(\text{crimen} == 1) = \frac{\exp(z_i)}{1 + \exp(z_i)}$$

$$z_i = \gamma_i + \gamma_2 \text{female}_i + \gamma_3 \text{Grupo} - \text{edad2} + \gamma_4 \text{Grupo} - \text{edad3}$$

$$+ \gamma_5 \text{Grupo} - \text{edad4} + \gamma_6 \text{Grupo} - \text{edad5} + \gamma_7 \text{SF} + \gamma_8 \text{Edad}$$

$$+ \gamma_9 \text{exporta} + \gamma_{10} \text{solefirms} + \gamma_{11} \text{manufacturing} + \gamma_{12} \text{commerce} + \gamma_{13} \text{services}$$

## Pregunta 2

- ▶ Ahora estimamos el modelo de regresión logístico:

$$\text{prob}(\text{crimen} == 1) = \frac{\exp(z_i)}{1 + \exp(z_i)}$$

$$z_i = \gamma_1 + \gamma_2 \text{female}_i + \gamma_3 \text{Grupo} - \text{edad2} + \gamma_4 \text{Grupo} - \text{edad3}$$

$$+ \gamma_5 \text{Grupo} - \text{edad4} + \gamma_6 \text{Grupo} - \text{edad5} + \gamma_7 \text{SF} + \gamma_8 \text{Edad}$$

$$+ \gamma_9 \text{exporta} + \gamma_{10} \text{solefirms} + \gamma_{11} \text{manufacturing} + \gamma_{12} \text{commerce} + \gamma_{13} \text{services}$$

- ▶ ¿Cuál es la propuesta de esta ecuación en este actual contexto?

## Pregunta 2

- ▶ Esta es la **ecuación de asignación de tratamiento** y esto es estimado usando un modelo logístico
- ▶ La variable dependiente en este caso (si la empresa sufre de crimen o no, `crime2`) representa los **tratados (treatment)**
- ▶ El modelo de regresión logisitica provee una estimación de **propensity scores**, el cuál predice las probabilidades de poder tene run crimen en este caso.
- ▶ La inclusión de variables debe influir el **tratamiento** y la variable **producto**, pero no influye en la variable **treatment**



## Pregunta 2

- ▶ Esta es la **ecuación de asignación de tratamiento** y esto es estimado usando un modelo logístico
- ▶ La variable dependiente en este caso (si la empresa sufre de crimen o no, `crime2`) representa los **tratados (treatment)**
- ▶ El modelo de regresión logisitica provee una estimación de **propensity scores**, el cuál predice las probabilidades de poder tene run crimen en este caso.
- ▶ La inclusión de variables debe influir el **tratamiento** y la variable **producto**, pero no influye en la variable **treatment**
- ▶ El supuesto crucial es el supuesto de independencia condicional (**CIA**, siglas en ingles).

## Pregunta 2



## Pregunta 2

- El supuesto implica que , dado ( o seleccionada) un conjunto de covariable observables que no se ven afectadas por el tratamiento ( selección de observables ), producto potenciales son independiente de la asignación del tratamiento.

## Pregunta 2

- ▶ El supuesto implica que , dado ( o seleccionada) un conjunto de covariable observables que no se ven afectadas por el tratamiento ( selección de observables ), producto potenciales son independiente de la asignación del tratamiento.
- ▶ Por lo tanto, la CIA implica que cualquier diferencia sistemática en los resultados entre los grupos *tratados* y *control* con los mismos valores de las covariables es atribuible al tratamiento.

## Pregunta 2

**Cuadro: Modelo Logistico**

	Model (1)	
==1 crime (Capital)		
gender condition	-0.19	(0.13)
30-36 years old (owned)	-0.32	(0.26)
36-46 years old (owned)	-0.39*	(0.20)
44-56 years old (owned)	-0.44**	(0.21)
More 55 years old (owned)	-0.53**	(0.23)
Access to capital credit	0.36***	(0.11)
Years firms	0.01	(0.01)
export status	0.26	(0.17)
Sole firms (propietario)	-0.06	(0.11)
sector2r4.2==Industrias manufactureras	-0.67***	(0.15)
sector2r4.2==Comercio	-0.25	(0.20)
sector2r4.2==Servicios	-0.22*	(0.13)
Constant	-2.09***	(0.25)
Observaciones	6115	
AIC	2976.7	

Fuente: INEI - ENE.

Elaboracion: Autor

\*\*\*, \*\*, \* denote statistical significance at the 1 %, 5 % and 10 % levels respectively for zero.

## Pregunta 2

### Cuadro: Modelo Logistico

	Model (1)	
==1 crime (Capital)		
gender condition	-0.19	(0.13)
30-36 years old (owned)	-0.32	(0.26)
36-46 years old (owned)	-0.39*	(0.20)
44-56 years old (owned)	-0.44**	(0.21)
More 55 years old (owned)	-0.53**	(0.23)
Access to capital credit	0.36***	(0.11)
Years firms	0.01	(0.01)
export status	0.26	(0.17)
Sole firms (propietario)	-0.06	(0.11)
sector2r4.2==Industrias manufactureras	-0.67***	(0.15)
sector2r4.2==Comercio	-0.25	(0.20)
sector2r4.2==Servicios	-0.22*	(0.13)
Constant	-2.09***	(0.25)
Observaciones	6115	
AIC	2976.7	

Fuente: INEI - ENE.

Elaboracion: Autor

\*\*\*, \*\*, \* denote statistical significance at the 1 %, 5 % and 10 % levels respectively for zero.

- Indicador de Performance *AIC*, para evaluar el desarrollo del modelo

## Pregunta 2

Cuadro: Modelo Logistico

	Model (1)	
==1 crime (Capital)		
gender condition	-0.19	(0.13)
30-36 years old (owned)	-0.32	(0.26)
36-46 years old (owned)	-0.39*	(0.20)
44-56 years old (owned)	-0.44**	(0.21)
More 55 years old (owned)	-0.53**	(0.23)
Access to capital credit	0.36***	(0.11)
Years firms	0.01	(0.01)
export status	0.26	(0.17)
Sole firms (propietario)	-0.06	(0.11)
sector2r4.2==Industrias manufactureras	-0.67***	(0.15)
sector2r4.2==Comercio	-0.25	(0.20)
sector2r4.2==Servicios	-0.22*	(0.13)
Constant	-2.09***	(0.25)
Observaciones	6115	
AIC	2976.7	

Fuente: INEI - ENE.

Elaboracion: Autor

\*\*\*, \*\*, \* denote statistical significance at the 1 %, 5 % and 10 % levels respectively for zero.

- Indicador de Performance *AIC*, para evaluar el desarrollo del modelo
- El Modelo Logistico permitira estimar el score de cada empresa dado el matching

## Pregunta 2

**Cuadro:** Modelo Logistico

	Model (1)	
==1 crime (Capital)		
gender condition	-0.19	(0.13)
30-36 years old (owned)	-0.32	(0.26)
36-46 years old (owned)	-0.39*	(0.20)
44-56 years old (owned)	-0.44**	(0.21)
More 55 years old (owned)	-0.53**	(0.23)
Access to capital credit	0.36***	(0.11)
Years firms	0.01	(0.01)
export status	0.26	(0.17)
Sole firms (propietario)	-0.06	(0.11)
sector2r4.2==Industrias manufactureras	-0.67***	(0.15)
sector2r4.2==Comercio	-0.25	(0.20)
sector2r4.2==Servicios	-0.22*	(0.13)
Constant	-2.09***	(0.25)
Observaciones	6115	
AIC	2976.7	

Fuente: INEI - ENE.

Elaboracion: Autor

\*\*\*, \*\*, \* denote statistical significance at the 1 %, 5 % and 10 % levels respectively for zero.

- ▶ Indicador de Performance *AIC*, para evaluar el desarrollo del modelo
- ▶ El Modelo Logistico permitira estimar el score de cada empresa dado el matching
- ▶ Se usa el modelo de regresión lineal logistica para calcular el **propensity score** ( la estimación de probabilidad )



## Pregunta 3



## Pregunta 3

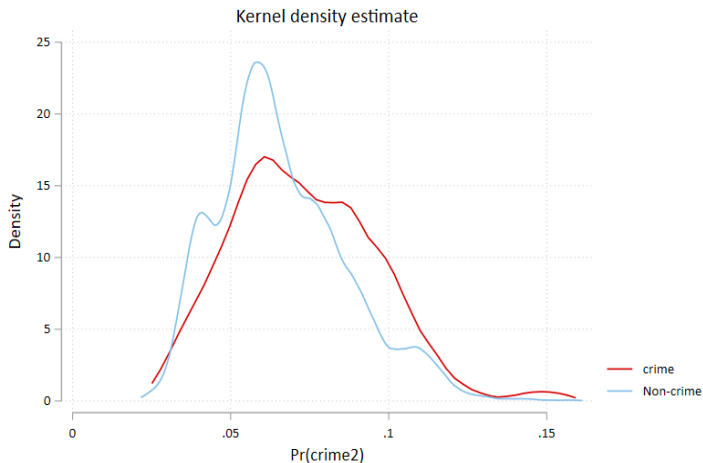
- ¿Cuántas observaciones están sobre **soporte común** en esta aplicación?

## Pregunta 3



## Pregunta 3

Figura: Densidad kernel de propensity score



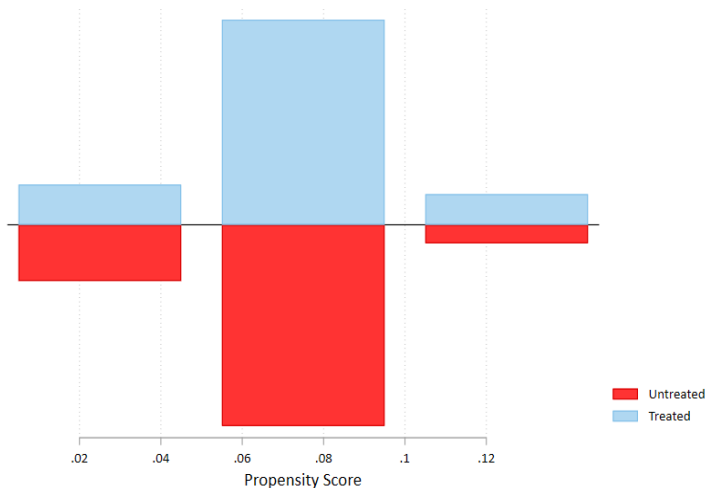
kernel = epanechnikov, bandwidth = 0.0061

## Pregunta 3



## Pregunta 3

**Figura:** Histograma de propensity score por treatment y control group



## Pregunta 3

```
. su phat if crime2==1
```

Variable	Obs	Mean	Std. Dev.	Min	Max
phat	407	.0732417	.0225737	.0311882	.1530884

```
.  
end of do-file
```

```
. do "C:\Users\edinson\AppData\Local\Temp\STD5ea4_000000.tmp"
```

```
. su phat if crime2==0
```

Variable	Obs	Mean	Std. Dev.	Min	Max
phat	5,708	.0660811	.0204855	.0215849	.1612774

## Pregunta 3

```
. su phat if crime2==1
```

Variable	Obs	Mean	Std. Dev.	Min	Max
phat	407	.0732417	.0225737	.0311882	.1530884

```
.  
end of do-file
```

```
. do "C:\Users\edinson\AppData\Local\Temp\STD5ea4_000000.tmp"
```

```
. su phat if crime2==0
```

Variable	Obs	Mean	Std. Dev.	Min	Max
phat	5,708	.0660811	.0204855	.0215849	.1612774

- La manera mas usual es chequear el **soporte comun** en los valores minimos y maximos del **propensity score** para ambas empresas que sufren de un crimen versus las que no.



## Pregunta 3

```
. su phat if crime2==1
```

Variable	Obs	Mean	Std. Dev.	Min	Max
phat	407	.0732417	.0225737	.0311882	.1530884

```
.  
end of do-file
```

```
. do "C:\Users\edinson\AppData\Local\Temp\STD5ea4_000000.tmp"
```

```
. su phat if crime2==0
```

Variable	Obs	Mean	Std. Dev.	Min	Max
phat	5,708	.0660811	.0204855	.0215849	.1612774

- ▶ La manera mas usual es chequear el **soporte comun** en los valores minimos y maximos del **propensity score** para ambas empresas que sufren de un crimen versus las que no.
- ▶ El rango en el **propensity score** para la empresas que sufre de crimen (treated), grupo **treatment** , es de **0.0311** para **0.1538**

## Pregunta 3

```
. su phat if crime2==1
```

Variable	Obs	Mean	Std. Dev.	Min	Max
phat	407	.0732417	.0225737	.0311882	.1530884

```
.  
end of do-file
```

```
. do "C:\Users\edinson\AppData\Local\Temp\STD5ea4_000000.tmp"
```

```
. su phat if crime2==0
```

Variable	Obs	Mean	Std. Dev.	Min	Max
phat	5,708	.0660811	.0204855	.0215849	.1612774

- ▶ La manera mas usual es chequear el **soporte comun** en los valores minimos y maximos del **propensity score** para ambas empresas que sufren de un crimen versus las que no.
- ▶ El rango en el **propensity score** para la empresas que sufre de crimen (treated), grupo **treatment** , es de **0.0311** para **0.1538**
- ▶ Algunos de los valores del **propensity score** para las empresas se encuentran fuera del rango predecido para grupo tratados (treatment)

## Pregunta 3



## Pregunta 3

- El **propensity score** para la empresas que no sufre de crimen (no crimen), grupo **control** tiene rango de **0.0215** para **0.1612**

## Pregunta 3

- ▶ El **propensity score** para la empresas que no sufre de crimen (no crimen), grupo **control** tiene rango de **0.0215** para **0.1612**
- ▶ Algunos de los valores de los **non crimen** tendrian que salir del **soporte comun**

## Pregunta 3

- ▶ El **propensity score** para la empresas que no sufre de crimen (no crimen), grupo **control** tiene rango de **0.0215** para **0.1612**
- ▶ Algunos de los valores de los **non crimen** tendrían que salir del **soporte comun**
- ▶ Estos son 26 non tea farmers fuera del **soporte comun** en este caso, y no deberían ser usados en el emparejamiento de *kernel*.

## Pregunta 4

**Figura:** Diagnostico de propiedad de balance

Variable	Mean		%bias	t-test		V(T)/ V(C)
	Treated	Control		t	p> t	
female	.18765	.20957	-5.4	-0.78	0.435	.
edad_group2	.09383	.08627	2.6	0.38	0.708	.
edad_group3	.30617	.30319	0.6	0.09	0.927	.
edad_group4	.36049	.36898	-1.8	-0.25	0.802	.
edad_group5	.23704	.25497	-4.2	-0.59	0.554	.
SF	.48148	.41698	13.0	1.85	0.065	.
Edad	12.457	12.336	2.0	0.28	0.779	1.03
exporta	.11111	.10467	2.1	0.29	0.768	.
solefirms	.38025	.39308	-2.6	-0.37	0.708	.
manufacturing	.25432	.30573	-11.3	-1.63	0.103	.
commerce	.0963	.09219	1.4	0.20	0.842	.
services	.37531	.38632	-2.3	-0.32	0.747	.

\* if variance ratio outside [0.82; 1.22]

## Pregunta 4

**Figura:** Diagnostico de propiedad de balance

► **Pregunta 4:** examine la propiedad de **balance** en esta aplicación

Variable	Mean		%bias	t-test		V(T)/ V(C)
	Treated	Control		t	p> t	
female	.18765	.20957	-5.4	-0.78	0.435	.
edad_group2	.09383	.08627	2.6	0.38	0.708	.
edad_group3	.30617	.30319	0.6	0.09	0.927	.
edad_group4	.36049	.36898	-1.8	-0.25	0.802	.
edad_group5	.23704	.25497	-4.2	-0.59	0.554	.
SF	.48148	.41698	13.0	1.85	0.065	.
Edad	12.457	12.336	2.0	0.28	0.779	1.03
exporta	.11111	.10467	2.1	0.29	0.768	.
solefirms	.38025	.39308	-2.6	-0.37	0.708	.
manufacturing	.25432	.30573	-11.3	-1.63	0.103	.
commerce	.0963	.09219	1.4	0.20	0.842	.
services	.37531	.38632	-2.3	-0.32	0.747	.

\* if variance ratio outside [0.82; 1.22]



## Pregunta 4

**Figura:** Diagnostico de propiedad de balance

Variable	Mean		%bias	t-test		V(T)/ V(C)
	Treated	Control		t	p> t	
female	.18765	.20957	-5.4	-0.78	0.435	.
edad_group2	.09383	.08627	2.6	0.38	0.708	.
edad_group3	.30617	.30319	0.6	0.09	0.927	.
edad_group4	.36049	.36898	-1.8	-0.25	0.802	.
edad_group5	.23704	.25497	-4.2	-0.59	0.554	.
SF	.48148	.41698	13.0	1.85	0.065	.
Edad	12.457	12.336	2.0	0.28	0.779	1.03
exporta	.11111	.10467	2.1	0.29	0.768	.
solefirms	.38025	.39308	-2.6	-0.37	0.708	.
manufacturing	.25432	.30573	-11.3	-1.63	0.103	.
commerce	.0963	.09219	1.4	0.20	0.842	.
services	.37531	.38632	-2.3	-0.32	0.747	.

\* if variance ratio outside [0.82; 1.22]

- **Pregunta 4:** examine la propiedad de **balance** en esta aplicación
- Una medida popular para medir la distancia en la distribución marginal de las covariables estandarizadas es el sesgo estandarizado (SB, standardized bias)

## Pregunta 4

**Figura:** Diagnostico de propiedad de balance

Variable	Mean		%bias	t-test		V(T)/ V(C)
	Treated	Control		t	p> t	
female	.18765	.20957	-5.4	-0.78	0.435	.
edad_group2	.09383	.08627	2.6	0.38	0.708	.
edad_group3	.30617	.30319	0.6	0.09	0.927	.
edad_group4	.36049	.36898	-1.8	-0.25	0.802	.
edad_group5	.23704	.25497	-4.2	-0.59	0.554	.
SF	.48148	.41698	13.0	1.85	0.065	.
Edad	12.457	12.336	2.0	0.28	0.779	1.03
exporta	.11111	.10467	2.1	0.29	0.768	.
solefirms	.38025	.39308	-2.6	-0.37	0.708	.
manufacturing	.25432	.30573	-11.3	-1.63	0.103	.
commerce	.0963	.09219	1.4	0.20	0.842	.
services	.37531	.38632	-2.3	-0.32	0.747	.

\* if variance ratio outside [0.82; 1.22]

- **Pregunta 4:** examine la propiedad de **balance** en esta aplicación
- Una medida popular para medir la distancia en la distribución marginal de las covariables estandarizadas es el sesgo estandarizado (SB, standardized bias)
- Como una **regla**, se puede esperar que el  $SB_A$  se puede ubicar entre 3 % y 5 % ( en terminos absolutos despues del emparejamiento exitoso.

## Pregunta 4

**Figura:** Diagnostico de propiedad de balance

Variable	Mean		%bias	t-test		V(T)/ V(C)
	Treated	Control		t	p> t	
female	.18765	.20957	-5.4	-0.78	0.435	.
edad_group2	.09383	.08627	2.6	0.38	0.708	.
edad_group3	.30617	.30319	0.6	0.09	0.927	.
edad_group4	.36049	.36898	-1.8	-0.25	0.802	.
edad_group5	.23704	.25497	-4.2	-0.59	0.554	.
SF	.48148	.41698	13.0	1.85	0.065	.
Edad	12.457	12.336	2.0	0.28	0.779	1.03
exporta	.11111	.10467	2.1	0.29	0.768	.
solefirms	.38025	.39308	-2.6	-0.37	0.708	.
manufacturing	.25432	.30573	-11.3	-1.63	0.103	.
commerce	.0963	.09219	1.4	0.20	0.842	.
services	.37531	.38632	-2.3	-0.32	0.747	.

\* if variance ratio outside [0.82; 1.22]

- **Pregunta 4:** examine la propiedad de **balance** en esta aplicación
- Una medida popular para medir la distancia en la distribución marginal de las covariables estandarizadas es el sesgo estandarizado (SB, standardized bias)
- Como una **regla**, se puede esperar que el  $SB_A$  se puede ubicar entre 3 % y 5 % ( en terminos absolutos despues del emparejamiento exitoso.
- El reporte de  $SB_B$  (standardized bias antes) reporta para la variable *land* es 47.3 %

## Pregunta 4

**Figura:** Diagnostico de propiedad de balance

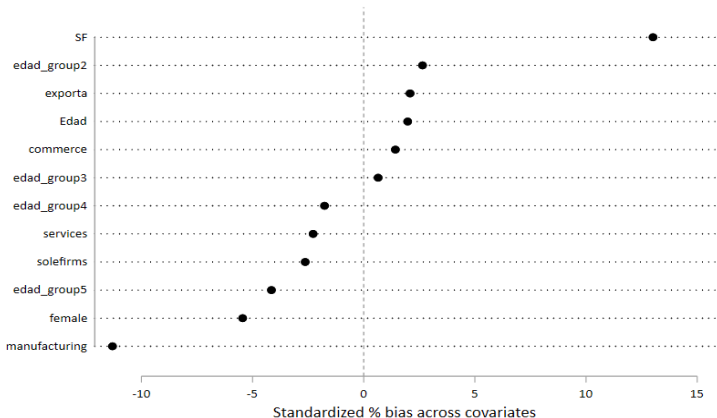
Variable	Mean		%bias	t-test		V(T)/ V(C)
	Treated	Control		t	p> t	
female	.18765	.20957	-5.4	-0.78	0.435	.
edad_group2	.09383	.08627	2.6	0.38	0.708	.
edad_group3	.30617	.30319	0.6	0.09	0.927	.
edad_group4	.36049	.36898	-1.8	-0.25	0.802	.
edad_group5	.23704	.25497	-4.2	-0.59	0.554	.
SF	.48148	.41698	13.0	1.85	0.065	.
Edad	12.457	12.336	2.0	0.28	0.779	1.03
exporta	.11111	.10467	2.1	0.29	0.768	.
solefirms	.38025	.39308	-2.6	-0.37	0.708	.
manufacturing	.25432	.30573	-11.3	-1.63	0.103	.
commerce	.0963	.09219	1.4	0.20	0.842	.
services	.37531	.38632	-2.3	-0.32	0.747	.

\* if variance ratio outside [0.82; 1.22]

- **Pregunta 4:** examine la propiedad de **balance** en esta aplicación
- Una medida popular para medir la distancia en la distribución marginal de las covariables estandarizadas es el sesgo estandarizado (SB, standardized bias)
- Como una **regla**, se puede esperar que el  $SB_A$  se puede ubicar entre 3 % y 5 % ( en terminos absolutos despues del emparejamiento exitoso.
- El reporte de  $SB_B$  (standardized bias antes) reporta para la variable *land* es 47.3 %
- y despues del matching, el  $SB_A$  se reduce a 5 %

## Pregunta 4

Figura: Standardized % Bias (despues del emparejamiento) por variables



## Pregunta 5

- Interprete precisamente el valor de ATT reportado para la variable resultado:  $\log(\text{productividad})$

## Pregunta 5



## Pregunta 5



- ▶ La expresión para el ATT basado en el emparejamiento de kernel ( $KM$ ) es:



## Pregunta 5

- ▶ La expresión para el ATT basado en el emparejamiento de kernel ( $KM$ ) es:
- ▶ donde hay  $N^T = 405$  empresas que sufrieron crimen, entonces esto es ATT es promedio sobre los 405 empresas

## Pregunta 5

- ▶ La expresión para el ATT basado en el emparejamiento de kernel ( $KM$ ) es:
- ▶ donde hay  $N^T = 405$  empresas que sufrieron crimen, entonces esto es ATT es promedio sobre los 405 empresas

$$ATT_{KM} = \frac{1}{N^T} \sum_{i \in T} \left( y_i^T - \sum_{i \in C} \pi_{ij} y_j^C \right)$$

## Pregunta 5

- ▶ La expresión para el ATT basado en el emparejamiento de kernel ( $KM$ ) es:
- ▶ donde hay  $N^T = 405$  empresas que sufrieron crimen, entonces esto es ATT es promedio sobre los 405 empresas

$$ATT_{KM} = \frac{1}{N^T} \sum_{i \in T} \left( y_i^T - \sum_{i \in C} \pi_{ij} y_j^C \right)$$

- ▶ Donde:

## Pregunta 5

- ▶ La expresión para el ATT basado en el emparejamiento de kernel ( $KM$ ) es:
- ▶ donde hay  $N^T = 405$  empresas que sufrieron crimen, entonces esto es ATT es promedio sobre los 405 empresas

$$ATT_{KM} = \frac{1}{N^T} \sum_{i \in T} \left( y_i^T - \sum_{i \in C} \pi_{ij} y_j^C \right)$$

- ▶ Donde:
  - ▶  $Y_i^T$ : el producto para treatment (sufrió crimen) unidad  $i$

## Pregunta 5

- ▶ La expresión para el ATT basado en el emparejamiento de kernel ( $KM$ ) es:
- ▶ donde hay  $N^T = 405$  empresas que sufrieron crimen, entonces esto es ATT es promedio sobre los 405 empresas

$$ATT_{KM} = \frac{1}{N^T} \sum_{i \in T} \left( y_i^T - \sum_{i \in C} \pi_{ij} y_j^C \right)$$

- ▶ Donde:
  - ▶  $Y_i^T$ : el producto para treatment (sufrió crimen) unidad  $i$
  - ▶  $Y_i^C$ : el producto para control (no sufrió de un crimen) unidad  $i$

## Pregunta 5

- ▶ La expresión para el ATT basado en el emparejamiento de kernel ( $KM$ ) es:
- ▶ donde hay  $N^T = 405$  empresas que sufrieron crimen, entonces esto es ATT es promedio sobre los 405 empresas

$$ATT_{KM} = \frac{1}{N^T} \sum_{i \in T} \left( y_i^T - \sum_{i \in C} \pi_{ij} y_j^C \right)$$

- ▶ Donde:
  - ▶  $Y_i^T$ : el producto para treatment (sufrió crimen) unidad  $i$
  - ▶  $Y_i^C$ : el producto para control (no sufrió de un crimen) unidad  $i$
  - ▶  $\pi_{ij}$ : el peso estandarizado calculado para cada  $i$  y  $j$  par usando los valores de Espanechnikov

## Pregunta 5

- ▶ La expresión para el ATT basado en el emparejamiento de kernel ( $KM$ ) es:
- ▶ donde hay  $N^T = 405$  empresas que sufrieron crimen, entonces esto es ATT es promedio sobre los 405 empresas

$$ATT_{KM} = \frac{1}{N^T} \sum_{i \in T} \left( y_i^T - \sum_{i \in C} \pi_{ij} y_j^C \right)$$

- ▶ Donde:
  - ▶  $Y_i^T$ : el producto para treatment (sufrió crimen) unidad  $i$
  - ▶  $Y_i^C$ : el producto para control (no sufrió de un crimen) unidad  $i$
  - ▶  $\pi_{ij}$ : el peso estandarizado calculado para cada  $i$  y  $j$  par usando los valores de Espanechnikov
  - ▶  $N^T$ : el tamaño muestral en el grupo de tratamiento  $i$

## Pregunta 5





## Pregunta 5



- ▶ Este es el resultado de *Linea Base* de ATT

## Pregunta 5

- ▶ Este es el resultado de *Linea Base* de ATT
- ▶ Se observa los datos de estandar error y el ratio-t para el ATT

**Figura:** STATA resultado del *PSMATCH2* para la linea base de las empresas que se vieron afectadas por el crimen

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
lnp1	Unmatched	8.4144238	8.43164064	-.017216839	.062006924	-0.28
	ATT	8.4144238	8.44523084	-.030807035	.059427279	-0.52

Note: S.E. does not take into account that the propensity score is estimated.