

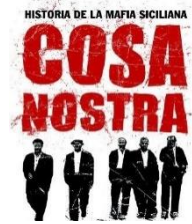
ACEMOGLU, DE FEO & DE LUCA (2020)

Weak States: Causes and Consequences of the Sicilian Mafia

[Daron Acemoglu](#), [Giuseppe De Feo](#), [Giacomo Davide De Luca](#)

The Review of Economic Studies, Volume 87, Issue 2, March 2020, Pages 537-581, <https://doi.org/10.1093/restud/rdz009>

Realizado por: Carolina Saavedra y Emily Saavedra



Acemoglu, De Feo y De Luca (2020) analizan la expansión de la mafia siciliana hacia finales del siglo XIX. Para ello, emplean dos variables explicativas principales: el nivel de precipitaciones en la primavera de 1893 y una variable dummy que indica la presencia del movimiento socialista conocido como *Peasant Fasci*.

La hipótesis central del estudio plantea que dicho movimiento logró consolidarse en aquellos distritos o regiones gravemente afectados por la sequía severa de ese año. Posteriormente, en un contexto marcado por una profunda crisis institucional, política, económica y social, el *Peasant Fasci* fue percibido como una amenaza por parte de las élites locales.

Esta percepción incentivó a terratenientes, administradores de propiedades agrícolas y actores políticos a fomentar la creación de agrupaciones mafiosas con el objetivo de contener la propagación del movimiento socialista.

Para desarrollar su análisis, los autores construyen una base de datos que incorpora las siguientes variables:

Tabla 1: Lista de variables empleadas en el estudio

Mafia1885	Distribución espacial de la Mafia reportado en Damiani (1885)
Mafia1890	Distribución espacial de la Mafia reportado en Cultrera (1890)
Agricola_rel	Parte de tierra total cultivada en 1853
Comune1853	Municipalidad como en 1853
Hhi1865	Índice de concentración de las elecciones parlamentarias de 1865 por municipalidad
Lnpop_1861	Log de población en 1861
Fasci	Dummy de la presencia de organización Fasci 1893-4
Relative rain	Lluvia relativa en la primavera de 1893 interpolada de las estaciones climáticas en 30 km

La variable *Fasci* se define como una variable dummy, mientras que *Mafia* es una variable categórica. A partir de ello, se procede a examinar empíricamente ambas dimensiones de la hipótesis planteada por los autores.

En primer lugar, se analiza la relación entre la presencia del movimiento *Peasant Fasci* y la ocurrencia de una sequía severa en 1893. La Tabla 2 presenta los promedios y desviaciones estándar del nivel de precipitaciones (*Relative Rain*) para los distritos según la presencia o ausencia de dicha organización:

Tabla 2: Nivel de precipitaciones y presencia de Peasant Fasci

Dummy que indica la presencia de Peasant Fasci	mean (relative rain)	sd (relative rain)
0	0.7511	0.2299
1	0.3710	0.2000

Como se observa, los distritos donde no se desarrolló el *Peasant Fasci* presentan, en promedio, mayores niveles de precipitaciones. Estos resultados son coherentes con la hipótesis inicial de los autores, quienes proponen el siguiente mecanismo causal:

↑ Sequía severa 1893 → ↑ Peasant Fasci 1893-94 → ↑ Mafia 1900.

Cabe señalar que la base de datos no incluye una variable explícita de sequía, sino una medida de precipitaciones relativa (*Relative Rain*). Por inferencia, niveles muy bajos de esta variable sugieren la ocurrencia de una sequía. En consecuencia, se esperaría una correlación negativa entre *Relative Rain* y *Fasci*. Al calcular dicha correlación, se obtiene un valor de -0.6136, lo que respalda empíricamente dicha relación.

En segundo lugar, se examina el impacto del *Peasant Fasci* sobre la aparición posterior de la mafia en los distritos, específicamente hacia 1900. Para ello, se estima la ecuación 1 mediante un modelo de probabilidad lineal (OLS), empleando errores estándar robustos.

Ecuación 01: Primera parte de la hipótesis

La especificación econométrica toma la siguiente forma:

$$Fasci_i = \alpha^{Fasci} \times relative\ rain_i^{1893} + X_i' \beta^{Fasci} + \varepsilon_i^{Fasci} \quad (1)$$

- *Fasci* representa la variable dummy que indica la presencia del movimiento socialista en el distrito.
- *Relative Rain* mide el nivel de precipitaciones en la primavera de 1893.
- *X* es un vector de variables de control.

La Tabla 3 presenta los resultados de la estimación de esta ecuación bajo distintas especificaciones.

Tabla 3: Ecuación 1 - Modelo MCO

Variable	Modelo sin variables de control	Modelo con variables de control			
	Modelo 1	Modelo 2	Modelo 3	Modelo 4	
	MPL_1	MPL_2	MPL_3	MPL_4	
Relative Rain	.2163***	-.8846***	-.9285***	-.9401***	
Lnpop_1861		.1019***	.05656**	.05848**	
Agricola_rel			.4238**	.4245**	
Mafia1885				-.0161	
N	297	297	297	297	
R2	0.08024429	0.5519709	0.56201887	0.56286889	
F	61.350218	108.17838	78.872974	60.134545	

Nota: *p<0.05; ** p<0.01; *** p<0.001

En la Tabla 3, el Modelo 1 muestra un coeficiente positivo para la variable Relative Rain, lo cual resulta contraintuitivo en relación con la hipótesis planteada por los autores. No obstante, en las demás especificaciones, los coeficientes estimados para dicha variable son negativos, tal como se anticipaba teóricamente.

A pesar de ello, los modelos estimados bajo MCO presentan un inconveniente metodológico relevante: predicen probabilidades negativas de ocurrencia del evento, lo cual es inconsistente con la naturaleza de los modelos de probabilidad. Este resultado sugiere la necesidad de emplear una metodología alternativa más adecuada para variables dependientes binarias.

Por ello, se opta por estimar un modelo Logit bivariado, controlando por heterocedasticidad, a través del método de Máxima Verosimilitud.



Tabla 4: Ecuación 1 - Modelo Logit Binomial

Variable	Modelo sin variables de control	Modelo con variables de control			
	Modelo 1	Modelo 2	Modelo 3	Modelo 4	
	Logit_1	Logit_2	Logit_3	Logit_4	
Relative Rain	-2.2397***	-6.2362***	-6.2807***	-6.4064***	
Lnpop_1861		.3027***	0.2359	0.2511	
Agricola_rel			0.6062	0.6278	
Mafia1885				-0.1227	
N	297	297	297	297	
x2	124.43447	85.707447	86.647567	86.423826	

Nota: *p<0.05; ** p<0.01; *** p<0.001

Posteriormente, se procede al análisis de los efectos marginales, dado que los coeficientes obtenidos mediante el modelo Logit no pueden interpretarse de manera directa en términos de probabilidades. Para tal propósito, se examinan dos especificaciones particulares: el Modelo 1 (Logit_1), que no incluye variables de control, y el Modelo 4 (Logit_4), que incorpora la totalidad de los controles contenidos en el vector *X*.

Tabla 5: Modelo Logit Binomial - Efectos marginales para la especificación 1 y 4

Variable	Em_Logit_1	Em_Logit_4
Relative Rain	-0.34741225	-0.76236593
Lnpop_1861		0.02987794
Agricola_rel		0.07470896
Mafia1885		-0.01460299

En la Tabla 5, los resultados del primer modelo muestran que un incremento en el nivel de precipitaciones durante la primavera de 1893 reduce la probabilidad de aparición del movimiento socialista Peasant Fasci en aproximadamente un 34%. Por su parte, el segundo modelo presenta un efecto aún más pronunciado: una disminución significativa en el nivel de precipitaciones —es decir, una sequía— se asocia con una reducción de alrededor del 76% en la probabilidad de que dicho movimiento se haya desarrollado.

Es importante señalar que estos efectos marginales han sido calculados individualmente para cada distrito y luego promediados. En consecuencia, los valores reportados corresponden a los efectos marginales promedio a nivel muestral.

Ecuación 02: Segunda parte de la hipótesis

Acemoglu, De Feo y De Luca (2020) consideran la variable *Mafia* como continua. No obstante, dado el carácter endógeno y jerárquico de esta variable, se opta en este análisis por tratarla como una variable categórica ordinal. En esta codificación, *Mafia* toma los

siguientes valores: 0 para ausencia de mafia, 1 para presencia mínima, 2 para nivel intermedio y 3 para nivel máximo de actividad mafiosa.

Este tratamiento se justifica en la medida en que las categorías presentan un orden natural y progresivo; es decir, para que un distrito pase de un nivel a otro, debe haber superado etapas previas de intensidad mafiosa. Por esta razón, se estima la Ecuación 2 utilizando un modelo logit multinomial ordenado.

$$Mafia_i = \alpha^{Mafia} \times Fasci_i + X_i' \beta^{Mafia} + \varepsilon_i^{Mafia} \quad (2)$$

En este modelo, la variable *Mafia* mide la intensidad mafiosa en el distrito *iii*; *Fasci* es una variable dummy que indica la presencia del movimiento *Peasant Fasci*; y *XXX* representa un vector de variables de control.

Se estiman cinco especificaciones distintas del modelo. La primera no incluye controles, mientras que las otras cuatro incorporan diferentes combinaciones de variables explicativas adicionales.

Tabla 6: Ecuación 2 - Logit Multinomial Ordenado

Variable	Modelo sin variables de control	Modelos con variables de control			
	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
	ologit_1	ologit_2	ologit_3	ologit_4	ologit_5
Fasci	1.7097262***	1.6606548***	1.6249189***	1.6659976***	0.62862078
Lnpop_1861		.45615836**	.45338574**	.39306246**	.46125309*
Agricola_rel			1.8965607	1.0206445	1.6722558
Mafia1885				.61332543***	.55347024***
Relative_rain				-3.0178469***	
/cut1	-.37161563*	3.4825132**	5.290602*	4.2404865	3.1450417
/cut2	.74055831***	4.6257186***	6.4374279**	5.4690612*	4.5113321
/cut3	2.035724***	5.9706823***	7.7831108**	6.9278297**	5.9598034*
Statistics					
N	273	273	273	273	245
Pseudo-R2	0.0706144	0.08598662	0.08732143	0.12764036	0.1650869
chi2	51.226992	59.06349	63.005388	83.118317	101.31624

Nota: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Los valores "cut" indican las estimaciones de los puntos de corte de la variable latente asociada a endógena del modelo.

Fuente: Elaboración propia

En la Tabla 6, los resultados estimados muestran que, tanto en el Modelo 1 como en el Modelo 2, todos los regresores son estadísticamente significativos. Sin embargo, los valores de pseudo- R^2 son relativamente bajos, lo cual sugiere una capacidad explicativa limitada. Por

otro lado, en los Modelos 4 y 5 se observan los valores más elevados de pseudo- R^2 y del estadístico X^2 , lo que indica un mejor ajuste general del modelo. No obstante, en la especificación 5, la variable *Fasci* pierde significancia estadística.

En consecuencia, se opta por centrar el análisis en el Modelo 4, ya que este ofrece evidencia empírica que respalda la hipótesis teórica planteada por los autores. En particular, se encuentra una asociación estadísticamente significativa al 1% entre la presencia del movimiento *Peasant Fasci* y la intensidad de la actividad mafiosa en Sicilia hacia finales del siglo XIX.

Dado el interés por estimar la probabilidad de que los distritos se encuentren en distintos niveles de intensidad mafiosa, se utiliza el comando `margins` para calcular los efectos marginales. En este ejercicio, los efectos marginales se calculan para cada distrito y luego se promedian. Los resultados se presentan en la Tabla 7.

Tabla 7: Logit Multinomial Ordenado-Efectos marginales para la especificación 4 "ologit_4"

Expression : Pr(Mafia==0), predict(outcome(0))						
	dy/dx	Delta-method Std. Err.	z	P>z	[95% Conf.	Interval]
Fasci	-0.2854	0.0402	-7.1000	0.0000	-0.3642	-0.2065
Inpop_1861	-0.0673	0.0225	-2.9900	0.0030	-0.1115	-0.0232
agricola_rel	-0.1748	0.3527	-0.5000	0.6200	-0.8661	0.5164
Mafia1885	-0.1051	0.0195	-5.4000	0.0000	-0.1432	-0.0669

Expression : Pr(Mafia==1), predict(outcome(1))						
	dy/dx	Delta-method Std. Err.	z	P>z	[95% Conf.	Interval]
Fasci	-0.0304	0.0125	-2.4300	0.0150	-0.0550	-0.0059
Inpop_1861	-0.0072	0.0042	-1.7000	0.0900	-0.0155	0.0011
agricola_rel	-0.0186	0.0395	-0.4700	0.6370	-0.0961	0.0588
Mafia1885	-0.0112	0.0051	-2.2100	0.0270	-0.0211	-0.0013

Expression : Pr(Mafia==2), predict(outcome(2))						
	dy/dx	Delta-method Std. Err.	z	P>z	[95% Conf.	Interval]
Fasci	0.0857	0.0163	5.2500	0.0000	0.0537	0.1177
Inpop_1861	0.0202	0.0074	2.7500	0.0060	0.0058	0.0347
agricola_rel	0.0525	0.1064	0.4900	0.6220	-0.1560	0.2610
Mafia1885	0.0316	0.0073	4.3100	0.0000	0.0172	0.0459

Expression : Pr(Mafia==3), predict(outcome(3))						
	dy/dx	Delta-method	z	P>z	[95% Conf.	Interval]
		Std. Err.				
Fasci	0.2301	0.0304	7.5800	0.0000	0.1706	0.2896
lnpop_1861	0.0543	0.0188	2.8900	0.0040	0.0175	0.0910
agricola_rel	0.1410	0.2853	0.4900	0.6210	-0.4182	0.7002
Mafia1885	0.0847	0.0155	5.4800	0.0000	0.0544	0.1150

En la Tabla 7, los valores reportados en la columna dy/dx representan los efectos marginales promedio, es decir, la probabilidad estimada de que los distritos se ubiquen en cada uno de los niveles de intensidad mafiosa definidos por la variable *Mafia*.

Para las categorías 0 (ausencia de mafia) y 1 (nivel mínimo), los coeficientes negativos y estadísticamente significativos de la variable *Fasci* indican que la presencia del movimiento *Peasant Fasci* reduce la probabilidad de que un distrito se encuentre en estos niveles bajos de actividad mafiosa. En contraste, para las categorías 2 (nivel intermedio) y 3 (nivel máximo), los coeficientes de *Fasci* son positivos y significativos al 1%, lo cual sugiere que los distritos que experimentaron dicho movimiento tienen una mayor probabilidad de alcanzar niveles más elevados de intensidad mafiosa.

A continuación, se realiza una comparación metodológica entre los resultados obtenidos al tratar la variable *Mafia* como una variable ordinal multinomial (modelo logit ordenado) y aquellos que se obtienen al considerarla como una variable continua.

En la sección anterior, se abordó la segunda parte de la hipótesis de los autores bajo el supuesto de que *Mafia* es una variable categórica ordenada. En lo que sigue, se estimará el mismo modelo asumiendo que *Mafia* es una variable continua.

Tabla 8: Ecuación 2 - Modelo de Regresión Lineal Simple

Variable	mco_1	mco_2	mco_3	mco_4	mco_5
Fasci	2.0792***	.9840***	.9415***	.9865***	.3796*
lnpop_1861		.1257***	.1423*	.1716**	.1841**
Agricola_rel			-0.3604	-0.4092	0.4214
Mafia1885				.3301***	.2745***
Relative Rain					-1.4773***
N	273	273	273	273	245
r ²	0.47615449	0.69349878	0.73038164	0.69410332	0.74622119
F	553.61585	374.02027	227.62373	251.13415	172.03631

Nota: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Fuente: Elaboración propia

Las estimaciones presentadas en la Tabla 8 presentan sesgos inherentes, ya que se basan en un modelo de regresión lineal simple estimado por Mínimos Cuadrados Ordinarios (OLS), el cual asume que la variable dependiente *Mafia* es continua. Bajo esta premisa, el

modelo interpreta que la variable puede adoptar cualquier valor en el intervalo real, incluyendo valores no enteros como 1.6 o 2.3, lo cual no refleja adecuadamente su naturaleza ordinal.

Como consecuencia, los coeficientes estimados reflejan un efecto promedio constante sobre todas las categorías de la variable, ignorando la heterogeneidad en los efectos que los regresores pueden tener sobre cada nivel específico de *Mafia*. Esta limitación es crítica, ya que, como se ha evidenciado previamente, dicha variable presenta un orden jerárquico natural que condiciona la progresión de los distritos a lo largo de niveles crecientes de intensidad mafiosa.

Conclusiones

El análisis confirma la importancia de emplear modelos econométricos acordes a la naturaleza de las variables y a la hipótesis teórica. El uso inicial de OLS permitió identificar correlaciones relevantes, pero sus limitaciones —como la predicción de probabilidades negativas— justificaron el paso a modelos logit. Estos ofrecieron resultados más consistentes y permitieron interpretar adecuadamente el vínculo entre sequía y aparición del *Peasant Fasci*.

Asimismo, al analizar la relación entre dicho movimiento y la intensidad mafiosa, se demostró que tratar la variable *Mafia* como ordinal (y no continua) permite capturar la heterogeneidad de los efectos según el nivel de actividad criminal. En conjunto, la estrategia metodológica adoptada refuerza la validez de los hallazgos y resalta cómo decisiones técnicas adecuadas enriquecen la interpretación de fenómenos históricos complejos.

