

## MÁLAGA

# Estimación de voto del estudio 3558 encuesta preelectoral de Andalucía. Abril de 2026

---

*Modelo bifactorial Inercia–Incertidumbre · output técnico*

Unidad de Ciencia Social Computacional

Gabinete de Presidencia

Centro de Investigaciones Sociológicas

### **Cómo citar este documento:**

Alaminos, Antonio y Alaminos-Fernández, Antonio Francisco (2026). *Estimación de voto de Málaga. Estudio encuesta preelectoral de Andalucía. Abril de 2026 (ECIS3558)*. Documentación técnica de la estimación electoral. Centro de Investigaciones Sociológicas. Madrid.

## Contenido

Marco metodológico: diseño de escenarios y naturaleza de los datos .....	4
Figura 1 .....	5
Figura 2 .....	8
Claves de lectura del informe .....	9
Alcance interpretativo, robustez y límites.....	11
Figura 3 .....	12
I. Análisis, medición y validación .....	13
1.1. Fundamentación metodológica del plano aritmético .....	13
1.2. Evaluación de Estimadores Centrales: Dialéctica entre Eficiencia y Robustez.....	13
1.3. Métricas de estabilidad y eficiencia estructural .....	14
1.4. Criterio de selección del estimador óptimo .....	16
1.5. Control de eficiencia estadística de los estimadores de localización .....	17
1.6. Inferencia mediante bootstrap estratificado por escenarios.....	21
Gráfico 1 — Mapa de decisión robusta.....	27
II. Diagnóstico validación estructural composicional (extensión CLR/ILR).....	28
Delimitación funcional entre el plano aritmético y el plano composicional.....	28
II.A. Análisis composicional exploratorio (geometría relacional global) .....	28
Fundamentación geométrica del análisis CLR.....	28
Biplot CLR.....	32
Gráfico 2 — Varianza explicada del PCA sobre coordenadas CLR (Scree plot) .....	34
Gráfico 3 — Visualización exploratoria del espacio relacional (Biplot composicional: PCA sobre CLR)...	35
II.B Análisis composicional confirmatorio (balances estructurales ILR).....	36
Diferenciación técnica: del rigor formal al diagnóstico estructural .....	36
Definición formal de los balances ILR por bloques.....	37
Balances ILR — Interpretación formal (razón geométrica) .....	38
Visualización exploratoria ILR (sin criterio estructural).....	41
Diagnóstico estructural ILR (cuadrantes e intensidad).....	42
Plano ILR por pares (diagnóstico estructural) .....	43
Diagrama Ternario: Equilibrio Relativo y Trayectoria de los Macro-bloques.....	45
Síntesis interpretativa de escenarios: variaciones y contribuciones estructurales .....	46
Validación bootstrap de los estimadores centrales bajo un esquema de re-muestreo de escenarios completos .....	49
II.C. Diagnóstico sistémico ILR ortonormal (base completa) .....	52
Reglas automáticas composicionales (CLR/ILR): validación estructural del estimado.....	54
Regla 1 — Estabilidad en % con reequilibrio composicional (CLR inestable).....	55
III. Gráficos Distribuciones bootstrap de estimadores de localización central.....	56
IV Bagplots bifactoriales entre partidos .....	59
Covariación bifactorial: PSOE vs PP .....	60
Covariación bifactorial: PSOE vs VOX .....	61
Covariación bifactorial: PSOE vs Adelante Andalucía.....	62
Covariación bifactorial: PSOE vs SALF.....	63
Covariación bifactorial: PSOE vs Por Andalucía .....	64
Covariación bifactorial: PP vs VOX.....	65
Covariación bifactorial: PP vs Adelante Andalucía .....	66
Covariación bifactorial: PP vs SALF .....	67
Covariación bifactorial: PP vs Por Andalucía .....	68
Covariación bifactorial: VOX vs Adelante Andalucía .....	69
Covariación bifactorial: VOX vs SALF .....	70
Covariación bifactorial: VOX vs Por Andalucía.....	71
Covariación bifactorial: Adelante Andalucía vs SALF.....	72
Covariación bifactorial: Adelante Andalucía vs Por Andalucía.....	73
Covariación bifactorial: SALF vs Por Andalucía.....	74
V. Anexo Técnico: Sistema Integrado de Estimación Robusta y Diagnóstico Estructural.....	75

Plano aritmético (criterio decisional) .....	75
Robustez estructural.....	75
Intensidad estructural.....	75
Plano relacional composicional .....	76
A. Plano Aritmético: Estimación Robusta del Centro.....	76
1. Media Aritmética .....	76
2. Mediana.....	76
3. Estimador de Huber (M-estimador) .....	77
B. Escalas Robustas .....	77
C. Índice de Robustez Penalizada (IRP*) .....	77
D. Diagnóstico de Forma y Segmentación.....	79
E. Eficiencia Bootstrap y Penalización .....	79
F. Regla de Selección del Estimador Óptimo .....	79
G. Propiedades del Sistema .....	80
H. Esquema Operativo del Sistema de Selección Óptima.....	80
Síntesis formal del sistema .....	81
Glosario de términos .....	83
V Referencias Metodológicas.....	85

## Marco metodológico: diseño de escenarios y naturaleza de los datos

**Introducción.** La demoscopia tradicional ha operado históricamente bajo la lógica de la 'foto fija': un intento de capturar la intención de voto en a) un instante preciso del tiempo, b) bajo el supuesto implícito de normalidad distributiva y c) en la hipótesis de estacionariedad relativa del fenómeno medido. Este enfoque, adecuado en contextos de competencia estable y baja volatilidad, resulta progresivamente limitado en sistemas electorales contemporáneos caracterizados por fragmentación, competencia multinivel y elevada sensibilidad a shocks exógenos. En este contexto, el Modelo Bifactorial Inercia–Incertidumbre surge como una respuesta metodológica que amplía el foco analítico incorporando, junto a la medición del apoyo electoral, una evaluación explícita de la estabilidad estructural de dicha medición.

Con tal fin, partiendo del análisis de los apoyos electorales que reciben los partidos políticos (descomposición univariable) se avanza hacia el análisis relacional sistémico mediante el espacio de Aitchison (composicional). Las validaciones cruzadas más importantes del diseño son cuatro. La primera contrasta estimadores dentro del plano aritmético mediante errores estándar y eficiencia bootstrap. La segunda compara estabilidad porcentual con estabilidad composicional mediante CLR. La tercera verifica en ILR si el centro elegido mantiene coherencia en los balances estructurales del sistema. La cuarta repite la comprobación mediante remuestreo de escenarios completos para ver si tanto el centro como la geometría relacional persisten.

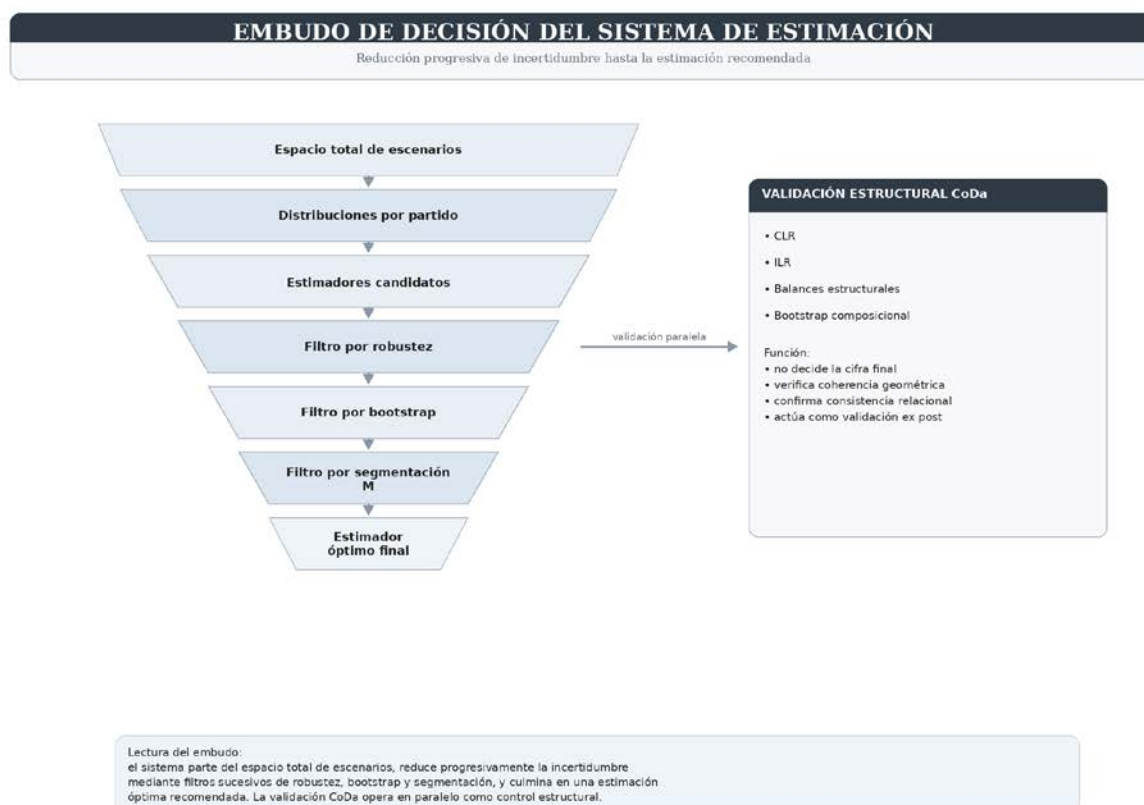
El modelo se fundamenta en un diseño de escenarios de estrés generados a partir de muestras aleatorias. La pregunta analítica deja de plantearse exclusivamente acerca de ¿cuántos votos tiene hoy un partido? y se reformula en términos de incertidumbre estructural: ¿cuál es el potencial electoral de un partido en el momento de la medición, considerando diferentes configuraciones teóricas del sistema? Esta reformulación amplía el foco para incluir, junto a la medición de la estimación puntual, un diagnóstico de la resiliencia e histéresis que muestra el apoyo electoral que recibe un partido frente a variaciones controladas en las condiciones estructurales que conforman los electorados (movilización y desmovilización, fidelidad y volatilidad). La arquitectura teórica del modelo, desarrollada en diversas publicaciones (Alaminos y Tezanos, 2029a, 2019b; Alaminos, 2021, 2022, 2024; Alaminos y Alaminos-Fernández, 2023a, 2023b, 2024), especifica un espacio de posibilidades proyectadas mediante la combinación sistemática de dos vectores críticos:

- El Factor Inercia: que representa la fuerza de retención y la lealtad histórica del voto.
- El Factor Incertidumbre: que representa la volatilidad inducida por eventos imprevistos, reconfiguraciones competitivas o cambios contextuales.

La interacción de ambos factores genera una matriz de escenarios en la que cada configuración (vector de voto) obliga al sistema multipartidista a comportarse bajo diferentes presiones estructurales. El objetivo no es producir una predicción determinista, sino evaluar la estabilidad relativa del sistema bajo condiciones de tensión controlada. En ese sentido establece un conjunto de estados potenciales, definidos a partir de una

composición de fuerzas (escenarios) condicionantes de las dinámicas que, en ausencia de impactos exógenos, dan forma al futuro. No es, en absoluto, una predicción del futuro y sí una medición y un diagnóstico del presente que le dará forma. Desde el punto de vista estadístico, el modelo adopta un enfoque de estimación robusta. Se calculan múltiples estimadores de localización (Media, Mediana, Huber con escala SD y MAD) evaluando mediante bootstrap su precisión relativa.

**Figura 1**



A partir del análisis de estabilidad de los coeficientes de localización central se establece una regla de selección automática que indica cuál es la estimación electoral más robusta. La regla de decisión automática selecciona el estimador óptimo maximizando la Eficiencia Penalizada:  $Eff^*(\theta) = Eff\_vs\_Mediana(\theta) \cdot (1 - M)$ . La eficiencia relativa  $Eff\_vs\_Mediana(\theta)$  se define como  $(SE(Mediana) / SE(\theta))^2$ , y la penalización estructural  $(1 - M)$  incorpora la intensidad de segmentación detectada en la distribución de escenarios. De forma complementaria, la Robustez Penalizada  $IRP^*$  evalúa la estabilidad intrínseca del centro porcentual ajustada por segmentación estructural:  $IRP^* = IRP \cdot (1 - M)$ , donde  $IRP = 1 / (1 + CV\_rob + S)$ . Esta separación permite distinguir conceptualmente entre precisión muestral, estabilidad del centro y estructura distributiva.

En consecuencia, el sistema puede expresarse formalmente como:

$$\text{diagnóstico\_estructural} = f(\text{Dip}, M, IRP^*)$$

selección\_del\_estimador =  $\text{argmax Eff}^*(\theta)$

Esta formulación separa explícitamente tres dimensiones analíticas: forma distributiva (Dip), segmentación estructural (M) y estabilidad del centro (IRP\*), evitando inferencias erróneas derivadas de confiar exclusivamente en contrastes paramétricos o en estadísticas de momento aisladas. El modelo resulta especialmente adecuado para tamaños muestrales moderados ( $n \approx 50-100$ ), donde la potencia de los contrastes formales puede ser insuficiente para detectar segmentaciones suaves o regímenes latentes. Un aspecto metodológico fundamental es la naturaleza composicional de los datos (suma constante = 100%).

En un sistema multipartidista, los componentes no son independientes: el crecimiento de un actor implica necesariamente la erosión relativa de otros. Ignorar esta restricción conduce a interpretaciones parciales o inconsistentes. Por ello, el informe incorpora un segundo nivel analítico basado en geometría composicional (Aitchison), mediante transformaciones log-ratio centradas (CLR) y balances ortonormales (ILR). Estas herramientas permiten distinguir entre estabilidad aritmética - constancia en porcentaje - y estabilidad relacional - constancia en las posiciones relativas dentro del sistema. La media composicional (Aitchison mean) se obtiene promediando en coordenadas CLR y aplicando la transformación inversa con cierre a suma 100. Los centros robustos composicionales se estiman mediante localización robusta por componente en el espacio CLR, aplicando recentrado a suma cero antes de invertir la transformación y cerrar nuevamente a 100. De este modo, el centro refleja proporciones relativas y no niveles absolutos, manteniendo la interdependencia entre componentes y evitando distorsiones producidas por outliers o reequilibrios compensatorios.

El análisis ILR complementa esta lectura identificando balances estructurales entre bloques de actores y evaluando intensidad radial y ocupación de cuadrantes. Ello permite distinguir entre tensiones compensatorias dentro de un mismo régimen y configuraciones estructurales alternativas bajo los escenarios considerados. Este informe se estructura en tres bloques analíticos complementarios. El primero determina el estimador puntual óptimo mediante criterios robustos y bootstrap. El segundo, valida la coherencia estructural y relacional del sistema bajo restricción composicional. Finalmente se procede a una segunda validación de los estimadores centrales y la estabilidad de los patrones observados aplicando procedimientos de re-muestreo. Consideradas conjuntamente, ambas aproximaciones permiten identificar tanto la mejor estimación central como la consistencia estructural del espacio competitivo que la sostiene.

El modelo bifactorial inercia-incertidumbre se concibe fundamentalmente como un instrumento de medición y diagnóstico. Su función no consiste únicamente en describir magnitudes, sino en proporcionar una base analítica para identificar la estructura relacional de los escenarios observados, estimar la intensidad de sus diferencias y reconocer los principales ejes de variación que organizan el espacio de competencia.

Desde esta perspectiva, el interés del análisis no reside solo en cuantificar apoyos, sino en detectar configuraciones, proximidades, contrastes y desplazamientos que permitan una lectura más precisa del estado del sistema.

Figura 2

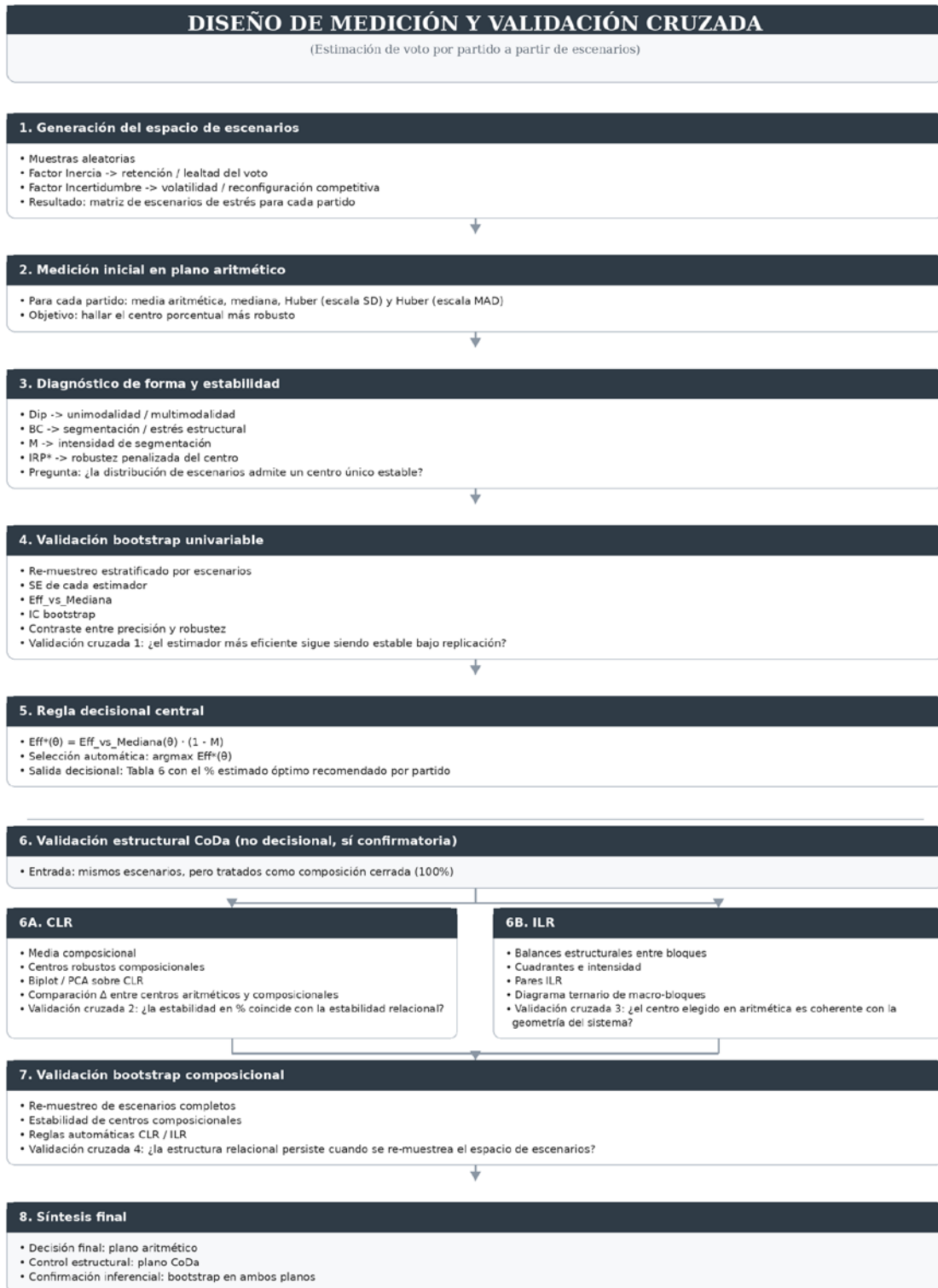


Figura. Esquema del proceso de medición robusta y validación cruzada del estimado de voto.

## Claves de lectura del informe

Este informe debe leerse desde una lógica composicional. Las magnitudes analizadas no representan valores aislados e independientes, sino partes de un conjunto cerrado cuya interpretación depende de las relaciones relativas entre componentes. Por ello, una variación en un elemento no debe entenderse solo en términos absolutos, sino como una alteración del equilibrio general del sistema. Esta premisa afecta tanto a la lectura de tablas y gráficos descriptivos como a la interpretación de los análisis geométricos posteriores.

La estructura del informe combina tres niveles de análisis complementarios. En primer lugar, un nivel descriptivo orientado a identificar diferencias visibles entre escenarios, configuraciones o estimaciones. En segundo lugar, un nivel estructural centrado en la posición relacional de las observaciones dentro del espacio composicional, mediante transformaciones log-ratio, planos ILR, representaciones ternarias y reducciones dimensionales. En tercer lugar, un nivel de validación destinado a evaluar la robustez de los estimadores centrales y la estabilidad de los patrones observados bajo procedimientos de re-muestreo.

Cada grupo de gráficos responde a una pregunta distinta. Los gráficos comparativos y los heatmaps permiten identificar la dirección y magnitud inmediata de los cambios. Los planos ILR, los simplex ternarios y los biplots sobre coordenadas CLR permiten interpretar esos mismos cambios como reconfiguraciones relacionales del sistema. Finalmente, las secciones de validación bootstrap añaden una capa de prudencia analítica al examinar hasta qué punto las posiciones centrales y los patrones detectados se mantienen estables cuando se altera el conjunto de observaciones considerado.

En consecuencia, el informe constituye una secuencia analítica escalonada. La lectura más inmediata de los cambios observables se complementa con una lectura estructural de su significado geométrico y con una evaluación posterior de su robustez. Esta organización permite distinguir entre variaciones visibles, pero analíticamente secundarias, desplazamientos composicionales relevantes y configuraciones cuya estabilidad aconseja una interpretación más o menos cautelosa.

El biplot composicional constituye un ejemplo particularmente útil de esta lógica de trabajo, en la medida en que traduce la información contenida en la matriz de datos a una representación sintética orientada al diagnóstico: permite situar los escenarios en un espacio común, identificar los componentes con mayor capacidad discriminante y observar la dirección general de las reconfiguraciones relativas. Su utilidad es, por tanto, eminentemente analítica, ya que facilita la interpretación estructurada de los datos y refuerza la capacidad del informe para medir y diagnosticar la dinámica interna del espacio político.

Como ejemplo, este biplot permite observar simultáneamente la posición relativa de los escenarios y la dirección de las principales tensiones composicionales del sistema. Los puntos resumen la proximidad o distancia entre configuraciones en el espacio reducido, mientras que las flechas indican qué componentes contribuyen con mayor intensidad a la variación relacional capturada por los dos primeros ejes.



Se trata, en definitiva, de una lectura geométrica y sintética del espacio bidimensional resumido por el PCA, útil para identificar afinidades, oposiciones y zonas de concentración estructural, pero no para establecer de manera mecánica relaciones causales ni supuestos de independencia estadística estricta.

### **Alcance interpretativo, robustez y límites**

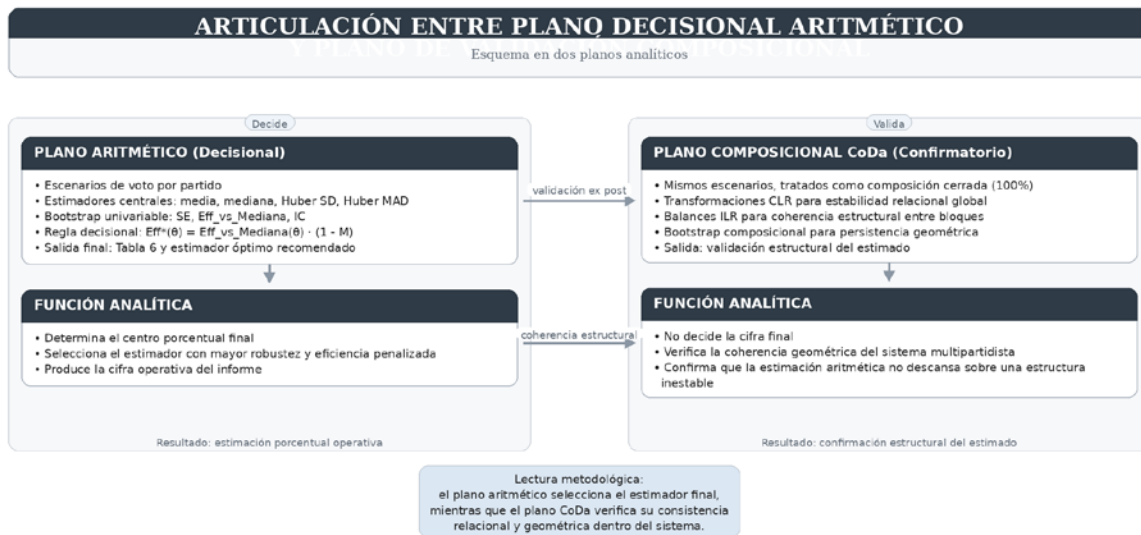
El valor principal de este enfoque reside en su capacidad para interpretar configuraciones complejas desde una perspectiva relacional y geoméricamente coherente con la naturaleza composicional de los datos. A diferencia de lecturas basadas únicamente en comparaciones directas de porcentajes o magnitudes aisladas, el análisis desarrollado aquí permite examinar la posición de cada observación dentro de un sistema de equilibrios, tensiones y desplazamientos entre componentes. Esta aproximación resulta especialmente útil cuando el interés analítico se centra en la estructura del conjunto y no solo en la evolución puntual de cada parte.

La incorporación de medidas robustas y procedimientos de validación bootstrap refuerza la consistencia del diagnóstico. Estos recursos permiten evaluar si las posiciones centrales y los patrones detectados dependen en exceso de observaciones concretas o si, por el contrario, muestran estabilidad bajo re-muestreo. En este sentido, la robustez no debe entenderse como una garantía absoluta, sino como una medida de confianza relativa en la persistencia de los resultados dentro del marco de escenarios y datos considerados.

Al mismo tiempo, conviene explicitar algunos límites del enfoque. En primer lugar, las representaciones geométricas y las distancias composicionales describen estructura y proximidad relativa, pero no establecen por sí mismas relaciones causales. En segundo lugar, una contribución alta a la distancia total indica relevancia en la reconfiguración del sistema, pero no equivale automáticamente a centralidad sustantiva en todos los planos de interpretación. En tercer lugar, la lectura de escenarios debe entenderse como una herramienta analítica para explorar configuraciones posibles o comparar estados del sistema, no como una predicción cerrada ni como una sustitución de otros enfoques explicativos.

Por ello, la interpretación final del informe debe apoyarse en una lectura prudente e integrada. Los resultados adquieren mayor valor cuando convergen varios indicadores: cambios descriptivos consistentes, desplazamientos visibles en el espacio composicional, contribuciones relevantes a la distancia total y estimadores relativamente estables bajo validación. Cuando esa convergencia es débil, el análisis sigue siendo útil, pero aconseja mayor cautela. Esta combinación entre capacidad descriptiva, lectura estructural y explicitación de límites constituye una de las principales fortalezas del enfoque.

Figura 3



En este informe, los bagplots se interpretan como representaciones bidimensionales del campo de probabilidad generado por los escenarios analizados. Su función no es predictiva, sino descriptivo-estructural: en lugar de estimar un único resultado futuro, delimitan el espacio de configuraciones compatibles con la distribución observada bajo las condiciones de inercia e incertidumbre del modelo. En este sentido, el bagplot actúa como una envolvente de densidad donde el 'bag' central concentra el núcleo de escenarios con mayor estabilidad, mientras que la región externa define la extensión del espacio de variación plausible. Esta geometría permite identificar simultáneamente la localización, dispersión y asimetría del sistema sin reducir la interpretación a una estimación puntual. Así, estos gráficos funcionan como herramientas de diagnóstico que describen la estructura probabilística analizada, enmarcando las zonas de proyección dentro de los límites definidos por el sistema de escenarios.

## I. Análisis, medición y validación

### 1.1. Fundamentación metodológica del plano aritmético

El plano aritmético aborda la estimación del centro como un problema de localización bajo incertidumbre estructural especificada por escenarios. Cada partido es tratado como una variable cuya distribución deriva del espacio bifactorial de inercia e incertidumbre. La selección del estimador central no se basa en un criterio único, sino en un enfoque comparativo entre media aritmética, mediana y estimador robusto de Huber. Esta comparación permite evaluar simultáneamente eficiencia, sensibilidad a extremos y estabilidad estructural. El objetivo no es describir la distribución, sino identificar el estimador que minimiza distorsión bajo variabilidad de escenarios, preservando coherencia con la lógica porcentual del dato.

### 1.2. Evaluación de Estimadores Centrales: Dialéctica entre Eficiencia y Robustez

La selección del estimador central óptimo dentro del plano aritmético no es una decisión meramente estadística, sino un compromiso epistemológico entre la representatividad y la estabilidad ante la incertidumbre. La media aritmética, si bien constituye el estimador de máxima eficiencia bajo condiciones de normalidad y simetría, presenta una vulnerabilidad crítica: su sensibilidad a los escenarios extremos. En sistemas multipartidistas sujetos a volatilidades abruptas, la media tiende a desplazarse artificialmente hacia los valores atípicos (outliers), lo que puede distorsionar el centro de gravedad real de la composición.

Como contraparte, la mediana ofrece el máximo nivel de robustez, poseyendo un punto de ruptura (breakdown point) del 50%. Su capacidad para ignorar las colas de la distribución la hace idónea para entornos de alta incertidumbre estructural; no obstante, esta resistencia conlleva un coste en términos de eficiencia. En distribuciones unimodales con varianza moderada, la mediana puede omitir información relevante de la masa de datos, resultando en una estimación menos precisa de la tendencia central comparada con estimadores de mayor orden. Para mitigar esta dicotomía, se introducen los estimadores M de Huber. Estos operan mediante una función de pérdida híbrida que se comporta de forma cuadrática cerca del centro (como la media) y lineal en las colas (como la mediana). Al implementar la variante Huber (escala SD), se prioriza la estabilidad basada en la desviación estándar, mientras que Huber (escala MAD) refuerza la resistencia mediante la desviación absoluta de la mediana. Este enfoque de ponderación adaptativa permite capturar la señal central del sistema sin que las fluctuaciones extremas de los escenarios simulados invaliden la métrica.

En consecuencia, la potencial divergencia entre estos cuatro estimadores en la Tabla 1 no expresa un error de medición, sino un diagnóstico de la asimetría estructural. Una brecha significativa entre la media y los estimadores de Huber indica la presencia de sesgos inducidos por la incertidumbre, donde la elección del estimador final -basada en la eficiencia penalizada ( $Eff^*$ )- determinará la fiabilidad del modelo composicional subsiguiente.

La Tabla 1 presenta los cuatro estimadores de localización considerados: Huber (escala SD), Huber (escala MAD), Media aritmética y Mediana. La media es eficiente bajo simetría, pero sensible a escenarios extremos. La mediana maximiza robustez, aunque puede perder eficiencia bajo distribuciones unimodales. El estimador de Huber equilibra ambas propiedades mediante ponderación adaptativa. Las diferencias observadas entre estimadores permiten evaluar las asimetrías inducidas por incertidumbre estructural.

**Tabla 1 — Estimadores centrales**

Partido	Huber (escala SD)	Huber (escala MAD)	Media aritmética	Mediana
PP	44.00	44.00	44.10	44.00
PSOE	25.70	25.80	25.60	26.10
VOX	9.70	9.70	9.80	9.30
Adelante Andalucía	7.00	7.00	7.00	7.00
Por Andalucía	5.90	5.80	5.80	5.90
Otro	3.80	3.80	3.80	3.80
SALF	3.40	3.40	3.40	3.40
En blanco	0.50	0.50	0.50	0.50

La coincidencia entre las estimaciones de Huber basadas en desviación típica y en MAD indica que las distribuciones presentan simetría y ausencia de valores extremos influyentes. Bajo estas condiciones, ambas medidas de escala convergen, generando ponderaciones equivalentes en el estimador robusto. Así, la proximidad entre media, mediana y estimadores robustos confirma la estabilidad probabilística del sistema y la ausencia de asimetrías estructurales relevantes en los escenarios. Por el contrario, diferencias sustantivas entre Huber-SD y Huber-MAD constituyen evidencia de heterocedasticidad, asimetría o presencia de outliers estructurales en la distribución.

En este contexto, la escala basada en desviación típica amplifica la influencia de colas largas, mientras que la basada en MAD reacciona con mayor robustez. La no convergencia entre ambos estimadores revela que el sistema no se encuentra en equilibrio distributivo y que la estimación puntual depende de la estructura interna de la variabilidad, lo que es consistente con un escenario de tensión o cambio de régimen.

### 1.3. Métricas de estabilidad y eficiencia estructural

La decisión del estimador central no se fundamenta únicamente en la magnitud puntual del centro, sino en métricas que evalúan su estabilidad frente a la variabilidad estructural inducida por los escenarios especificados. En este marco, la comparación entre media, mediana y estimadores robustos (Huber-SD y Huber-MAD) se articula con los indicadores de eficiencia (Eff) y eficiencia penalizada (Eff\*), que combinan precisión y estabilidad distributiva. Eff mide la relación entre dispersión y coherencia interna del estimador seleccionado, permitiendo identificar cuál ofrece menor varianza efectiva bajo el conjunto de escenarios bootstrap. Eff\*, por su parte, introduce una penalización adicional cuando se detectan asimetrías, divergencias entre escalas robustas o separación significativa entre media y mediana, incorporando así una dimensión de robustez estructural.

La elección final no privilegia simplemente el valor central más frecuente o cercano a la media, sino aquel que maximiza la eficiencia penalizada, es decir, el que mantiene mayor estabilidad geométrica y probabilística ante perturbaciones del modelo. Con tal fin la selección del estimador central se basa en un criterio de eficiencia comparada donde  $Eff$  evalúa la relación entre precisión y dispersión bajo supuestos de estabilidad distributiva, mientras que  $Eff^*$  incorpora penalizaciones asociadas a inestabilidad estructural (divergencia entre Huber-SD y Huber-MAD, asimetría, multimodalidad o separación entre media y mediana).

De este modo, la decisión final responde a un principio de mínima vulnerabilidad ante perturbaciones del espacio probabilístico bootstrap, priorizando estimadores cuya posición se mantiene consistente tanto en términos porcentuales como en términos robustos. Este procedimiento asegura que la estimación recomendada responda a regularidades persistentes del sistema y no a configuraciones contingentes de escenarios específicos. En resumen, la decisión del estimador central no se fundamenta únicamente en la magnitud del centro, sino en métricas que evalúan su estabilidad bajo variabilidad de escenarios:

- $Eff^*$  (eficiencia penalizada) integra precisión bootstrap y estabilidad estructural, penalizando sensibilidad excesiva a escenarios extremos.
- IRP\* (Índice de Robustez Penalizada) mide la resiliencia del estimador frente a la variabilidad inducida por el espacio de escenarios, tomando valores entre 0 y 1.
- $M$  (intensidad estructural) captura el grado de dispersión generado por inercia e incertidumbre. Valores bajos indican Segmentación; valores elevados sugieren sensibilidad estructural.

**Tabla 2 — Evaluación de Robustez**

Partido	Volatilidad	Sensibilidad a extremos	Índice de robustez (IRP)	Segmentación interna (BC, si Dip=Unimodal)
PP	0.03	0.00	0.97	Segmentación alta
PSOE	0.05	0.01	0.94	Segmentación baja
VOX	0.10	0.02	0.90	Segmentación baja
Adelante Andalucía	0.05	0.01	0.94	Segmentación alta
Por Andalucía	0.17	0.00	0.85	Segmentación baja
Otro	0.18	0.00	0.84	Segmentación alta
SALF	0.09	0.01	0.91	
En blanco	0.23	0.00	0.81	Segmentación baja

La combinación de  $Eff^*$ ,  $IRP^*$  y  $M$  permite comparar estimadores no solo por su localización puntual, sino por su comportamiento bajo el conjunto completo de escenarios. El estimador óptimo será aquel que maximice estabilidad y eficiencia penalizada, preservando coherencia porcentual. Cuando  $Eff$  y  $Eff^*$  convergen en el mismo estimador y las métricas robustas coinciden, se interpreta un régimen de estabilidad distributiva. Por el contrario, divergencias entre  $Eff$  y  $Eff^*$  señalan tensiones internas o posibles cambios de régimen probabilístico, requiriendo una lectura más cautelosa del estimador central recomendado.

#### 1.4. Criterio de selección del estimador óptimo

La Tabla 3 sintetiza el proceso decisional del plano aritmético. La selección final del estimador central se realiza exclusivamente en este espacio, sin intervención de transformaciones composicionales. El criterio decisional prioriza el estimador con mayor  $Eff^*$ , condicionado por niveles adecuados de  $IRP^*$  y baja intensidad estructural ( $M$ ).

**Tabla 3 — Forma distribucional y diagnóstico**

Partido	Dip (p-valor)	Modalidad (Dip)	BC Sarle (consenso/estrés)	Score multimodalidad (M) desde BC	Robustez ajustada (IRP*)	Intensidad estructural (M)
PP	0.72	Unimodal	0.47	0.00	0.97	Baja fragmentación latente
PSOE	0.67	Unimodal	0.64	0.19	0.76	Baja fragmentación latente
VOX	0.58	Unimodal	0.70	0.33	0.60	Fragmentación incipiente
Adelante Andalucía	0.59	Unimodal	0.46	0.00	0.94	Baja fragmentación latente
Por Andalucía	0.14	Unimodal	0.57	0.04	0.82	Baja fragmentación latente
Otro	0.83	Unimodal	0.42	0.00	0.84	Baja fragmentación latente
SALF	0.02	Multimodal	0.55	0.00	0.91	Segmentación leve
En blanco	0.80	Unimodal	0.60	0.11	0.73	Baja fragmentación latente

La prueba de Dip y el Coeficiente de Bimodality (BC) evalúan la morfología de la distribución del partido a través de los escenarios. El Dip (p-valor) determina la presencia de multimodalidad ( $p < 0,05$  implica rechazo de unimodalidad), mientras que el BC no decide modalidad, sino que actúa como indicador de segmentación y estrés estructural. En contextos unimodales, valores elevados de BC pueden señalar aplanamiento o proximidad a transición ( $\approx 0,555$ ). El Score M, derivado y normalizado a partir del BC, cuantifica la intensidad de segmentación entre escenarios y ajusta la robustez mediante  $IRP^* = IRP \cdot (1 - M)$ . El IRP mide la resiliencia del partido frente a la variabilidad de escenarios; un IRP\* alto indica posición estructuralmente estable y poco dependiente de las fluctuaciones del sistema.

### 1.5. Control de eficiencia estadística de los estimadores de localización

La Tabla 1.4a. presenta los errores estándar (SE) obtenidos mediante bootstrap para cada estimador de localización central aplicado al conjunto de escenarios, así como la eficiencia relativa respecto a la mediana ( $Eff_{vs\ Mediana}$ ). El SE mide la variabilidad del estimador entre los escenarios especificados teóricamente; valores más bajos indican mayor estabilidad muestral. La eficiencia relativa se define como  $Eff_{vs\ Mediana}(\theta) = \left(\frac{SE(Mediana)}{SE(\theta)}\right)^2$ . Valores superiores a 1 indican que el estimador  $\theta$  es más eficiente que la mediana en términos de precisión bootstrap; valores inferiores a 1 indican menor eficiencia. Es importante subrayar que esta eficiencia es puramente aritmética y no incorpora todavía penalización estructural. La selección final del estimador óptimo se realizará posteriormente mediante  $Eff^*(\theta) = Eff_{vs\ Mediana}(\theta) \cdot (1 - M)$ , donde M captura la intensidad de segmentación estructural.

**Tabla 4a — Errores estándar (SE) y eficiencia relativa vs mediana (Eff)**

Partido	SE Huber(SD)	SE Huber(MAD)	SE Media	SE Mediana	Eff vs Med (Huber SD)	Eff vs Med (Huber MAD)	Eff vs Med (Media)
PP	0,20	0,22	0,19	0,19	0,89	0,78	1,07
PSOE	0,20	0,24	0,20	0,24	1,50	1,02	1,48
VOX	0,16	0,21	0,15	0,19	1,43	0,87	1,72
Adelante Andalucía	0,06	0,06	0,06	0,07	1,11	1,12	1,19
Por Andalucía	0,11	0,11	0,10	0,10	0,85	0,85	1,04
Otro	0,10	0,10	0,09	0,10	0,99	1,02	1,21
SALF	0,03	0,03	0,03	0,03	0,73	0,70	0,73
En blanco	0,01	0,01	0,01	0,02	2,97	2,78	3,63

En términos sustantivos, los SE más bajos indican las mayores eficiencias relativas bajo las perturbaciones inducidas por los escenarios. Los estimadores Huber (SD y MAD) habitualmente presentan un comportamiento intermedio: mejoran la robustez frente a posibles valores extremos, pero pueden mostrar ligera pérdida de eficiencia respecto a la media cuando la distribución no exhibe fuerte segmentación.

En partidos con tamaños muestrales reducidos o con baja variabilidad estructural pueden observarse eficiencias cercanas a 1 o valores no definidos cuando los errores estándar son prácticamente nulos. Estos casos reflejan estabilidad extrema o ausencia de variabilidad detectable en el conjunto de escenarios. La tabla debe interpretarse, por tanto, como un diagnóstico preliminar de precisión muestral entre estimadores. La decisión definitiva dependerá de la integración con la intensidad estructural ( $M$ ) y la robustez penalizada ( $IRP^*$ ), evitando seleccionar estimadores únicamente por eficiencia aritmética cuando existe segmentación distributiva relevante.

La Tabla 4b integra la precisión muestral y la estructura distributiva mediante la eficiencia penalizada  $Eff^*(\theta) = Eff_{vs\ Mediana}(\theta) \cdot (1 - M)$ . A diferencia de la Tabla 4a, aquí la eficiencia no se evalúa únicamente por variabilidad bootstrap, sino que incorpora la intensidad de segmentación estructural  $M$ . La columna 'Mejor (Eff)' identifica el estimador con mayor eficiencia aritmética pura (sin penalización). La columna 'Mejor (Eff\*)' refleja la decisión final del sistema, una vez incorporada la penalización estructural. Esta distinción es fundamental: un estimador puede ser el más preciso en términos bootstrap, pero no necesariamente el más adecuado cuando existe segmentación distributiva relevante.

**Tabla 4b — Eficiencia penalizada (Eff\*) y mejor estimador (Best)**

Partido	Eff* Huber(SD)	Eff* Huber(MAD)	Eff* Media	Mejor (Eff)	Mejor (Eff*)
PP	0,89	0,78	1,07	Media	Media
PSOE	1,21	0,83	1,20	Huber_mean	Huber_mean
VOX	0,96	0,58	1,15	Media	Media
Adelante Andalucía	1,11	1,12	1,19	Media	Media
Por Andalucía	0,82	0,82	1,00	Media	Media
Otro	0,99	1,02	1,21	Media	Media
SALF	0,73	0,70	0,73	Mediana	Mediana
En blanco	2,65	2,48	3,24	Media	Media

Los casos marcados con asterisco (\*) corresponden a distribuciones bootstrap prácticamente puntuales ( $SE \approx 0$ ). En estos escenarios no existe variabilidad detectable entre escenarios y cualquier estimador converge, por lo que la elección no tiene implicaciones sustantivas. Los casos con doble asterisco (\*\*) indican multimodalidad extrema ( $M > 0,95$ ). En estos supuestos la penalización estructural hace que  $Eff^*$  tienda a cero para todos los estimadores univariados. Esto refleja que la distribución no puede representarse adecuadamente mediante un único centro. Se recomienda en tales casos la modelización mediante mezclas (por ejemplo, Gaussian Mixture Models) en lugar de un estimador puntual.

En los principales partidos estatales frecuentemente la media mantiene la mayor eficiencia incluso tras penalización estructural, lo que indica ausencia de segmentación intensa y estabilidad del régimen composicional bajo los escenarios simulados. En partidos pequeños habitualmente la mediana resulta preferible tras considerar la estructura distributiva. Ello sugiere presencia de asimetrías o pequeñas segmentaciones que reducen la conveniencia de la media como estimador robusto.

La tabla formaliza la regla decisional del modelo:  $\hat{\theta} = \operatorname{argmax} Eff^*(\theta)$ . La selección no responde únicamente a precisión estadística, sino a la coherencia entre estabilidad muestral y estructura distributiva, garantizando que el centro recomendado sea consistente con la morfología real del espacio de escenarios. En conjunto las tablas 4a y 4b muestran la integración precisión-estructura donde la Tabla 1.5. proporciona el diagnóstico de precisión muestral entre estimadores, midiendo la variabilidad bootstrap de cada centro bajo el conjunto de escenarios mientras que la Tabla 4b incorpora dicha información dentro de una arquitectura decisional que penaliza explícitamente la segmentación estructural mediante el factor  $(1 - M)$ . Ambas tablas deben leerse conjuntamente.

Así, en primer lugar, cuando un estimador presenta simultáneamente menor SE y mayor  $Eff\_vs\_Mediana$  en la Tabla 4a, y además mantiene la mayor  $Eff^*$  en la Tabla 4b, existe coherencia plena entre precisión muestral y estabilidad estructural. En estos casos la decisión es robusta en ambos planos. En segundo lugar, cuando el estimador más eficiente en términos aritméticos pierde posición tras aplicar la penalización estructural, la diferencia entre 'Mejor ( $Eff$ )' y 'Mejor ( $Eff^*$ )' revela la presencia de segmentación distributiva suficiente como para desaconsejar la selección basada exclusivamente en precisión.

Este mecanismo evita una selección basada únicamente en el error estándar y garantiza coherencia entre forma distributiva y centro recomendado. La penalización estructural no modifica los SE observados, sino que ajusta la decisión en función de la morfología real del espacio de escenarios. Los casos con SE prácticamente nulo reflejan estabilidad extrema entre escenarios. En estos supuestos la eficiencia relativa carece de relevancia sustantiva, dado que todos los estimadores convergen hacia el mismo valor. Por el contrario, cuando  $M$  se aproxima a 1 (multimodalidad o segmentación extrema), la eficiencia penalizada  $Eff^*$  tiende a cero para todos los estimadores univariados. Esta situación indica que el problema no es de

selección de estimador, sino de representación inadecuada mediante un único centro, justificando el recurso a modelos de mezcla.

En términos formales, la arquitectura decisional puede resumirse como:

- 1) Diagnóstico de precisión: SE y Eff\_vs\_Mediana.
- 2) Diagnóstico estructural: M (intensidad de segmentación).
- 3) Decisión integrada:  $\hat{\theta} = \operatorname{argmax} \operatorname{Eff}^*(\theta)$ .

Esta secuencia asegura separabilidad analítica entre variabilidad muestral y morfología distributiva, evitando inferencias sesgadas bajo heterogeneidad estructural. En consecuencia, es relevante advertir que las tablas 4a y 4b no muestran únicamente una comparación técnica de estimadores, sino un sistema integrado de decisión robusta que armoniza precisión estadística y coherencia estructural.

## 1.6. Inferencia mediante bootstrap estratificado por escenarios

Las distribuciones bootstrap permiten evaluar empíricamente la estabilidad de cada estimador bajo replicación vectorial del espacio de escenarios. Este procedimiento evita supuestos paramétricos de normalidad y garantiza consistencia inferencial bajo posibles asimetrías inducidas por incertidumbre.

**Tabla 5 — Control de validación: estimación Bootstrap univariable**

Partido	Media (IC 95%)	Huber SD (IC 95%)	Huber MAD (IC 95%)
PP	44,1 [43,7, 44,5]	44,0 [43,6, 44,4]	44,0 [43,6, 44,4]
PSOE	25,6 [25,2, 26,0]	25,7 [25,3, 26,1]	25,8 [25,3, 26,3]
VOX	9,8 [9,6, 10,1]	9,7 [9,5, 10,1]	9,7 [9,3, 10,1]
Adelante Andalucía	7,0 [6,9, 7,2]	7,0 [6,9, 7,1]	7,0 [6,9, 7,1]
Por Andalucía	5,8 [5,6, 6,0]	5,9 [5,6, 6,1]	5,8 [5,6, 6,0]
Otro	3,8 [3,6, 4,0]	3,8 [3,6, 4,0]	3,8 [3,6, 4,0]
SALF	3,4 [3,3, 3,5]	3,4 [3,3, 3,5]	3,4 [3,3, 3,5]
En blanco	0,4 [0,4, 0,4]	0,4 [0,4, 0,4]	0,4 [0,4, 0,4]

La inferencia mediante bootstrap estratificado por escenarios aproxima la distribución muestral empírica de cada estimador de localización  $\hat{\theta}$  sin imponer supuestos paramétricos sobre la forma distributiva subyacente.

Sea  $\hat{\theta}_b$  la replicación bootstrap del estimador en la iteración  $b$ , con  $b = 1, \dots, B$ . A partir de ese conjunto de repeticiones pueden definirse la varianza empírica, el sesgo bootstrap y los intervalos de confianza percentílicos.

$$Var_{boot}(\hat{\theta})$$

$$Bias_{boot}(\hat{\theta})$$

$$IC_{boot}^{0.95}$$

Valores reducidos de  $Bias_{boot}(\hat{\theta})$  indican estabilidad del centro bajo perturbación controlada y ausencia de efectos sistemáticos derivados de la estructura del espacio proyectado. La consistencia bajo re-muestreo se evalúa mediante la convergencia entre estimadores alternativos y la estabilidad de  $Var_{boot}(\hat{\theta})$  frente a variaciones en  $B$ .

Cuando la dispersión bootstrap permanece acotada y la distancia entre estimadores robustos y media aritmética es pequeña, puede inferirse que la estructura distributiva no presenta colas pesadas ni multimodalidad significativa en el régimen considerado.

Desde la perspectiva del modelo bifactorial, la estabilidad inferencial observada se integra en la función de eficiencia penalizada  $Eff^*(\hat{\theta}) = Eff_{vs\ Mediana}(\hat{\theta}) \cdot (1 - M)$ , donde la penalización estructural  $M$  captura la segmentación efectiva del espacio de escenarios. Una  $Var_{boot}(\hat{\theta})$  reducida y un  $Bias_{boot}(\hat{\theta})$  próximo a cero limitan la magnitud de  $M$  en tanto que la segmentación empírica es débil, reforzando la coherencia entre precisión muestral y estabilidad estructural. En conjunto, el análisis bootstrap formaliza tres

dimensiones complementarias: dispersión empírica  $Var_{boot}(\hat{\theta})$ , sesgo estructural  $Bias_{boot}(\hat{\theta})$  y cobertura  $IC_{boot}^{0.95}$ . Esta tríada permite validar la robustez del centro estimado bajo replicación estratificada y proporciona una base inferencial sólida para la posterior selección automática del estimador óptimo mediante criterios de eficiencia penalizada.

Los intervalos de confianza bootstrap  $IC_{boot}^{0.95} = [Q_{0.025}(\hat{\theta}_b), Q_{0.975}(\hat{\theta}_b)]$  se interpretan en términos de replicación empírica del estimador bajo el diseño estratificado, garantizando validez inferencial incluso en presencia de asimetrías o restricciones composicionales.

Desde la perspectiva del modelo bifactorial, la estabilidad inferencial observada se integra en la función de eficiencia penalizada.

$$Eff^*(\hat{\theta}) = Eff_{vs\ Mediana}(\hat{\theta}) \cdot (1 - M)$$

En esta expresión,  $M$  captura la segmentación efectiva del espacio de escenarios. Una varianza bootstrap reducida y un sesgo próximo a cero limitan la magnitud de  $M$  cuando la segmentación empírica es débil, reforzando la coherencia entre precisión muestral y estabilidad estructural.

En conjunto, el análisis bootstrap formaliza tres dimensiones complementarias: dispersión empírica, sesgo estructural y cobertura inferencial. Esta tríada permite validar la robustez del centro estimado bajo replicación estratificada y proporciona una base sólida para la selección automática del estimador óptimo.

La Tabla 6 sintetiza la decisión automática de selección del estimador óptimo para cada componente del sistema multipartidista.

La selección se realiza conforme a la función de eficiencia penalizada y a los indicadores de robustez estructural  $IRP^*$  y  $M$ .

$$Eff^*(\hat{\theta}) = Eff_{vs\ Mediana}(\hat{\theta}) \cdot (1 - M)$$

$$\hat{\theta} = argmax_{\theta \in \theta} Eff^*(\theta)$$

El porcentaje estimado corresponde al valor normalizado recomendado tras aplicar esa regla, garantizando coherencia entre precisión muestral y estabilidad estructural. La columna  $Eff^*$  refleja la eficiencia relativa frente a la mediana ajustada por penalización estructural, mientras que  $IRP^*$  captura la estabilidad intrínseca del centro bajo segmentación.

**Tabla 6 — Estimador óptimo según robustez**

Partido	% estimado (óptimo recomendado normalizado)	Estimador óptimo recomendado	Justificación (valor original, Eff*, IRP*, M)
PP	44,1	Media	Valor=44.093; Eff*=1.07, IRP*=0.97, M=0.00
PSOE	25,7	Huber_mean	Valor=25.747; Eff*=1.21, IRP*=0.76, M=0.19
VOX	9,8	Media	Valor=9.849; Eff*=1.15, IRP*=0.60, M=0.33
Adelante Andalucía	7,0	Media	Valor=7.049; Eff*=1.19, IRP*=0.94, M=0.00
Por Andalucía	5,8	Media	Valor=5.813; Eff*=1.00, IRP*=0.82, M=0.04
Otro	3,8	Media	Valor=3.777; Eff*=1.21, IRP*=0.84, M=0.00
SALF	3,4	Mediana	Valor=3.382; Eff*=1.00, IRP*=0.91, M=0.00
En blanco	0,4	Media	Valor=0.416; Eff*=3.24, IRP*=0.73, M=0.11

**Tabla 6a — Tratamiento sugerido para estadios de Multimodalidad Fuerte ( $M > 0,80$ )**

Partido	Score M	Modalidad	Estimador actual	Diagnóstico estructural	Recomendación avanzada
---------	---------	-----------	------------------	-------------------------	------------------------

Nota técnica. Para valores de  $M$  superiores a 0,80 el sistema detecta segmentación estructural del electorado. El uso de Huber mejora la robustez frente a extremos, pero no modeliza la coexistencia de subpoblaciones latentes. En estos casos se recomienda estimar un Gaussian Mixture Model (GMM) para identificar centros múltiples y sus probabilidades asociadas.

La Tabla 6a identifica configuraciones en las que el indicador de segmentación estructural  $M$  supera el umbral crítico  $M > 0,80$ , lo que implica evidencia empírica de multimodalidad fuerte en la distribución bootstrap de escenarios. En tales casos, la hipótesis de unicidad del centro  $\theta$  como estimador representativo del régimen

competitivo resulta estadísticamente inconsistente, dado que la densidad empírica sugiere la coexistencia de subpoblaciones latentes. Formalmente, la distribución observada puede modelizarse como una mezcla finita de  $K$  componentes, definida como  $f(x) = \sum_k \pi_k \phi(x | \mu_k, \Sigma_k)$ , donde  $\pi_k \geq 0$ ,  $\sum_k \pi_k = 1$  representan los pesos relativos de cada régimen latente, y  $\phi(x | \mu_k, \Sigma_k)$  denota la función de densidad normal multivariante con media  $\mu_k$  y matriz de covarianza  $\Sigma_k$ .

Bajo este esquema, el centro global deja de tener interpretación estructural unívoca y pasa a depender de la combinación convexa de componentes. En presencia de segmentación elevada, la penalización estructural incorporada en  $Eff^*(\hat{\theta}) = Eff_{vs\ Mediana}(\hat{\theta}) \cdot (1 - M)$  reduce automáticamente la eficiencia efectiva del estimador univariante, reflejando que  $Var_{boot}(\hat{\theta})$  puede permanecer acotada aun cuando la distribución subyacente sea multimodal. De este modo, el parámetro  $M$  actúa como corrector de especificación frente a errores derivados de asumir homogeneidad distributiva.

El Modelado de Mezclas Gaussianas (GMM) permite estimar los parámetros  $\pi_k, \mu_k, \Sigma_k$  mediante máxima verosimilitud, habitualmente implementada a través del algoritmo EM. Esta aproximación desplaza la inferencia desde la estimación de un único  $\theta$  hacia la identificación de centros múltiples  $\mu_k$  y sus probabilidades asociadas  $\pi_k$ , proporcionando una representación estructural coherente del espacio de escenarios cuando la incertidumbre induce bifurcaciones del régimen competitivo. En términos metodológicos, la incorporación de modelos de mezcla distingue entre dispersión muestral -capturada por  $Var_{boot}(\hat{\theta})$  y  $IC_{boot}^{0.95}$  - y heterogeneidad estructural genuina -capturada por  $M$ -, evitando confundir ruido inferencial con fragmentación latente. Así, el procedimiento mantiene coherencia con la arquitectura del Modelo Bifactorial al reconocer que, bajo incertidumbre elevada, el sistema puede transitar entre múltiples configuraciones estables en lugar de converger hacia un único equilibrio central.

La Tabla 7 presenta el diagnóstico estructural integrado de cada componente del sistema, combinando evidencia de modalidad (test de Dip), estructura modal (coeficiente de bimodalidad, BC), intensidad de segmentación ( $M$ ) y robustez penalizada (IRP\*). Este bloque sintetiza la dimensión morfológica de la distribución bootstrap de escenarios, permitiendo distinguir entre regímenes unimodales estables y configuraciones con fragmentación latente.

El estadístico Dip evalúa formalmente la hipótesis de unimodalidad frente a alternativas multimodales, mientras que el coeficiente BC proporciona una medida complementaria de forma distributiva. La intensidad  $M$  cuantifica la magnitud efectiva de segmentación estructural, incorporándose como penalización en la función de eficiencia  $Eff^*(\hat{\theta})$ . Por su parte, IRP\* resume la estabilidad del centro estimado ajustada por fragmentación. Este diagnóstico no sustituye la estimación puntual óptima, sino que la contextualiza estructuralmente. Mientras la selección del estimador responde a criterios de eficiencia penalizada, la presente

tabla permite evaluar si el centro estimado representa un régimen único estable o si se encuentra condicionado por tensiones internas o bifurcaciones latentes del espacio de escenarios.

**Tabla 7 — Diagnóstico por partido (Dip + Segmentación + Intensidad + Robustez)**

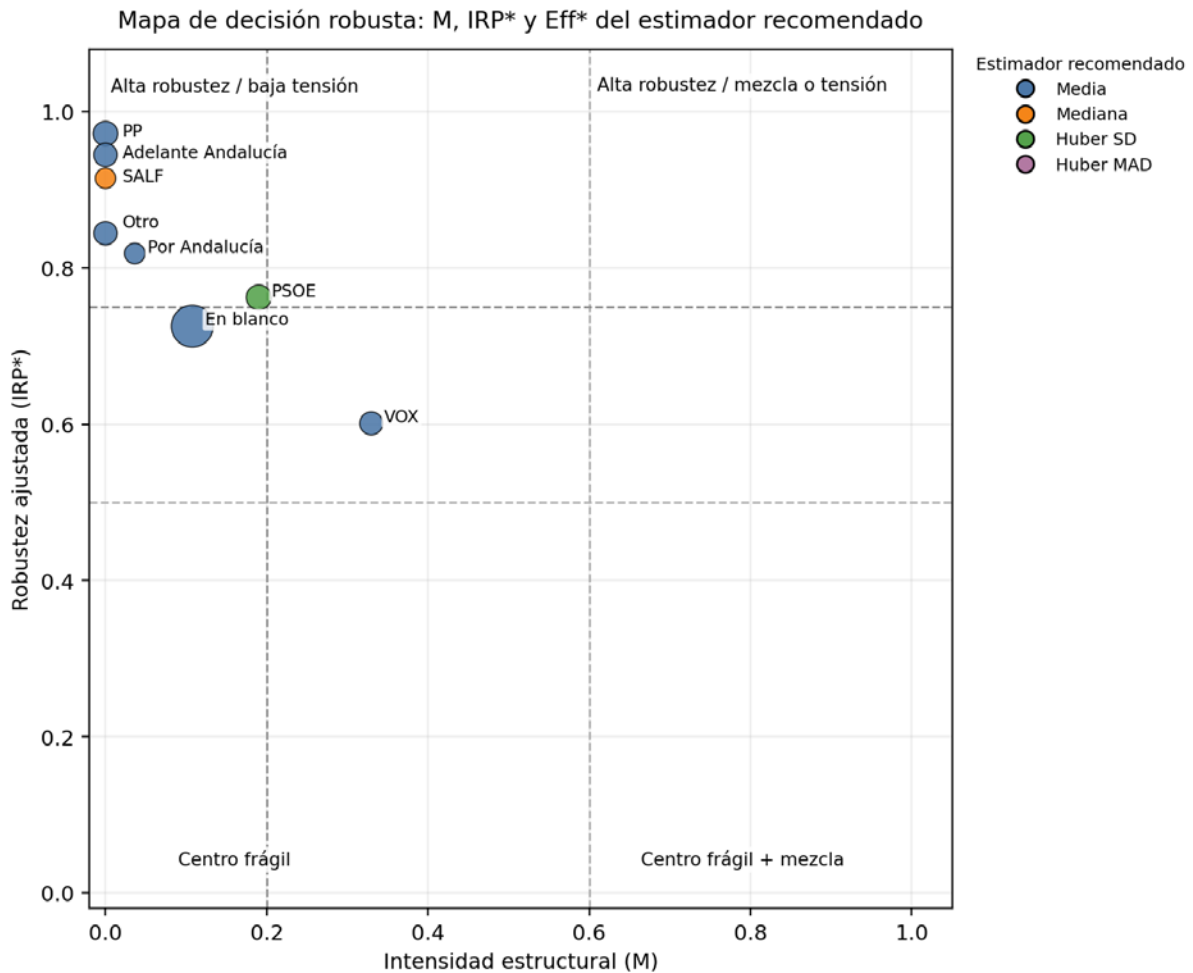
Partido	% estimado	Modalidad (Dip)	Estructura modal (BC)	Intensidad (M)	Robustez (IRP*)
PP	44,1	Unimodal	Segmentación alta	Baja fragmentación latente	Alta
PSOE	25,7	Unimodal	Segmentación baja	Baja fragmentación latente	Alta
VOX	9,8	Unimodal	Segmentación baja	Fragmentación incipiente	Alta
Adelante Andalucía	7,0	Unimodal	Segmentación alta	Baja fragmentación latente	Alta
Por Andalucía	5,8	Unimodal	Segmentación baja	Baja fragmentación latente	Alta
Otro	3,8	Unimodal	Segmentación alta	Baja fragmentación latente	Alta
SALF	3,4	Multimodal	n/a	Segmentación leve	Alta
En blanco	0,4	Unimodal	Segmentación baja	Baja fragmentación latente	Alta

Los resultados del diagnóstico estructural permiten diferenciar conceptualmente entre dispersión muestral, segmentación efectiva y estabilidad del régimen competitivo. Una configuración unimodal con IRP\* elevado sugiere coherencia distributiva y estabilidad bajo replicación bootstrap. En contraste, la presencia de multimodalidad o valores elevados de M indica que el espacio de escenarios puede contener subestructuras latentes, aun cuando la estimación puntual permanezca formalmente definida.

Es importante subrayar que el plano aritmético constituye el nivel decisional del modelo. La selección del estimador central se realiza exclusivamente en este espacio, garantizando coherencia porcentual y eficiencia estadística. Los análisis estructurales y composicionales operan como validación ex post, permitiendo verificar que la estimación óptima no se apoya en configuraciones distributivas inestables o en regímenes mixtos no detectados por estadísticos clásicos. En consecuencia, la Tabla 7 cumple una función de control estructural dentro de la arquitectura bifactorial: integra forma distributiva (Dip), intensidad de segmentación (M) y robustez ajustada (IRP\*) para asegurar que la estimación central seleccionada refleja un equilibrio competitivo consistente bajo incertidumbre controlada.

El gráfico sintetiza en un único plano la lógica decisional del sistema. El eje horizontal representa la intensidad estructural ( $M$ ), el eje vertical la robustez ajustada ( $IRP^*$ ), el tamaño de cada punto refleja la eficiencia penalizada ( $Eff^*$ ) del estimador finalmente recomendado y el color identifica el tipo de estimador seleccionado. Esta visualización permite distinguir, de forma inmediata, entre configuraciones estables, centros frágiles y casos en los que la mezcla estructural aconseja mayor cautela.

### Gráfico 1 — Mapa de decisión robusta



Las líneas discontinuas señalan umbrales de referencia del propio sistema. En términos sustantivos, valores bajos de  $M$  y altos de  $IRP^*$  indican centros consistentes y poco tensionados, mientras que incrementos de  $M$  o descensos de  $IRP^*$  sugieren segmentación latente, mezcla o menor estabilidad del centro estimado.

## II. Diagnóstico validación estructural composicional (extensión CLR/ILR)

Este análisis añade un plano composicional al diagnóstico principal. Los porcentajes están sujetos a restricción de suma constante; por ello, la variación en un actor siempre implica compensación en otros. Las transformaciones log-ratio de Aitchison permiten evaluar estabilidad en términos relativos (log-cocientes) y separar volatilidad aritmética de reestructuración composicional. Valida la coherencia estructural de la estimación mediante técnicas composicionales (CLR/ILR), que respetan la restricción de suma constante y la naturaleza relativa del sistema multipartidista.

### Delimitación funcional entre el plano aritmético y el plano composicional

En el marco del análisis univariable, las reglas aritméticas implementadas operan bajo un principio de optimización estadística cuyo objetivo unívoco es la determinación del estimador central de mayor robustez.

Dicha selección se fundamenta en la convergencia entre la eficiencia penalizada  $Eff^*$  y la intensidad estructural  $M$ . Esta fase constituye un proceso analítico autónomo, cuya validez interna es independiente de las transformaciones de carácter composicional.

Por el contrario, el bloque de transformaciones log-ratio (CLR/ILR) desempeña una función estrictamente validativa respecto a la coherencia estructural del sistema multipartidista. Al operar bajo la restricción de suma constante (cierre unitario), este plano no interviene en la elección del estimador central, sino que garantiza que la estimación sea geoméricamente consistente dentro del simplex.

En consecuencia, se establece una dicotomía funcional clara: el indicador  $IRP^*$  dictamina la estabilidad del centro de gravedad en el espacio euclídeo aritmético, mientras que las coordenadas CLR e ILR describen la estabilidad relacional y la arquitectura interna de los componentes. Dado que estos indicadores operan en espacios métricos de distinta naturaleza, sus interpretaciones son complementarias pero no intercambiables.

### II.A. Análisis composicional exploratorio (geometría relacional global)

Se traslada el análisis desde la medición en el plano porcentual (aritmético) al plano composicional para integrar la restricción de suma constante (cada escenario suma 100). Se emplean transformaciones log-ratio (Aitchison) con la finalidad de evaluar estabilidad en términos de relaciones proporcionales y no de niveles absolutos.

#### Fundamentación geométrica del análisis CLR

Sea  $\mathbf{x} = (x_1, \dots, x_D)$  una composición con  $x_i > 0$  y cierre unitario. Internamente el sistema trabaja con cierre a suma 100, pero la interpretación geométrica se mantiene.

La transformación centered log-ratio (CLR) proyecta cada composición al espacio euclídeo y permite expresar cada componente en términos relativos respecto a la media geométrica del conjunto.

$$clr(x_i) = \ln\left(\frac{x_i}{g(x)}\right)$$

En coordenadas CLR, cada componente expresa su ventaja o desventaja relativa respecto al promedio geométrico del sistema. Las coordenadas CLR viven en un subespacio de dimensión  $D - 1$  con restricción de suma cero.

Por ello, cuando se estiman centros por componente en CLR, el programa aplica un recentrado posterior a suma cero antes de invertir la transformación; este paso es necesario para que la inversión CLR sea geoméricamente coherente.

La tabla 8 muestra los centros composicionales obtenidos de dos formas coherentes con Aitchison:

- (i) la media composicional (Aitchison mean), que equivale a promediar en CLR y transformar de vuelta aplicando la inversa CLR y un cierre final a suma 100; y
- (ii) dos centros robustos Huber (escala SD y MAD), calculados por componente en CLR, posteriormente recentrados a suma cero, invertidos a simplex y cerrados a 100.

**Tabla 8 — Centros composicionales (media geométrica cerrada y estimadores robustos)**

Partido	Media_comp	HuberSD_comp	HuberMAD_comp
PP	44.2	44.2	44.2
PSOE	25.6	25.7	25.7
VOX	9.8	9.7	9.7
Adelante Andalucía	7.1	7.0	7.0
Por Andalucía	5.8	5.8	5.8
Otro	3.7	3.8	3.8
SALF	3.4	3.4	3.4
En blanco	0.4	0.4	0.4

Los centros composicionales se estiman como medias geométricas cerradas asegurando que el centro refleje relaciones proporcionales y no niveles absolutos. La proximidad entre media geométrica y centros robustos indica estabilidad global del sistema en términos relativos.

Las diferencias entre estimadores aritméticos y composicionales cuantifican el efecto de la restricción de suma constante. Valores próximos a cero indican que la estimación porcentual no está distorsionada por reequilibrios internos del sistema. Diferencias sistemáticas señalarían que la estabilidad aritmética podría estar sosteniéndose sobre compensaciones proporcionales ocultas.

La tabla 9 muestra la diferencia  $\Delta$  entre centros composicionales y estimadores aritméticos del bloque I. Estas diferencias no sustituyen la decisión del estimador óptimo (que se realiza en el plano aritmético), sino que informan si la solución porcentual es coherente con el centro geométrico del sistema bajo clausura. Este contraste permite distinguir entre estabilidad aritmética genuina y estabilidad aparente derivada de redistribuciones internas entre actores.

**Tabla 9 — Diferencias respecto a estimadores aritméticos ( $\Delta$  composicional)**

Partido	$\Delta(\text{Media\_comp} - \text{Media\_arit})$	$\Delta(\text{HuberSD\_comp} - \text{Huber\_SD\_arit})$	$\Delta(\text{HuberMAD\_comp} - \text{Huber\_MAD\_arit})$
PP	0.1	0.2	0.2
PSOE	0.0	0.0	-0.1
VOX	0.0	0.0	0.0
Adelante Andalucía	0.1	0.0	0.0
Por Andalucía	0.0	-0.1	0.0
Otro	-0.1	0.0	0.0
SALF	0.0	0.0	0.0
En blanco	-0.1	-0.1	-0.1

La matriz de variaciones constituye el pilar fundamental para la caracterización de la estabilidad interna en sistemas multipartidistas. A diferencia de las matrices de covarianza tradicionales, que se ven distorsionadas por el sesgo de 'correlación espuria' inherente a los datos de suma constante, la matriz de variaciones  $V$  opera sobre la estructura logarítmica de las proporciones. Formalmente, se define mediante el elemento  $v_{ij} = \text{Var}[\ln(x_i/x_j)]$ , el cual cuantifica la dispersión de la relación proporcional entre el partido  $x_i$  y el partido  $x_j$  a través de los diferentes escenarios temporales o espaciales estudiados.

Desde una perspectiva geométrica, esta matriz captura la métrica intrínseca del espacio de Aitchison. Cada entrada  $v_{ij}$  actúa como un indicador de codependencia estructural: valores cercanos a cero en  $v_{ij}$  denotan una estabilidad casi perfecta en la relación de fuerzas entre dos formaciones políticas, sugiriendo que sus bases electorales fluctúan de manera estrictamente proporcional. Por el contrario, valores elevados en  $v_{ij}$  revelan una alta volatilidad relativa, identificando los puntos de mayor tensión y reconfiguración dentro del sistema de partidos. La implementación de la matriz  $T_{ij} = \text{Var}[\ln(x_i/x_j)]$  permite trascender las limitaciones del simplex, transformando un espacio restringido en un entorno euclídeo vectorial mediante el uso de log-ratios. Este enfoque garantiza la invarianza por escala y la coherencia subcomposicional, propiedades esenciales para asegurar que las inferencias sobre la estructura multipartidista no dependan de la inclusión o exclusión arbitraria de fuerzas minoritarias. En última instancia, la matriz de variaciones no solo describe la dispersión, sino que revela la arquitectura latente de la competencia política.

**Tabla 10 — Matriz de variaciones log-ratio (estructura de covariación relativa):  $\text{Var}[\log(x_i/x_j)]$**

	PP	PSOE	VOX	Adelante Andalucía	Por Andalucía	Otro	SALF	En blanco
PP	0,000	0,007	0,017	0,006	0,022	0,050	0,010	0,036
PSOE	0,007	0,000	0,028	0,015	0,012	0,056	0,016	0,057
VOX	0,017	0,028	0,000	0,006	0,064	0,022	0,003	0,010
Adelante Andalucía	0,006	0,015	0,006	0,000	0,045	0,038	0,005	0,018
Por Andalucía	0,022	0,012	0,064	0,045	0,000	0,097	0,043	0,104
Otro	0,050	0,056	0,022	0,038	0,097	0,000	0,031	0,033
SALF	0,010	0,016	0,003	0,005	0,043	0,031	0,000	0,017
En blanco	0,036	0,057	0,010	0,018	0,104	0,033	0,017	0,000

La interpretación es directa: valores cercanos a cero indican alta estabilidad relativa entre dos partidos, mientras que valores elevados reflejan inestabilidad o sensibilidad estructural al escenario. Una variación muy baja implica que la razón entre ambos partidos permanece prácticamente constante en todos los escenarios, lo que sugiere comportamiento paralelo o electorados estructuralmente vinculados. Desde una lectura política son electorados en relación estable: si uno crece o decrece, el otro lo hace en proporción relativa similar.

Por el contrario, valores elevados indican que la relación entre ambos partidos es volátil o incierta. Identifican sensibilidad sistémica como expresión de interacciones entre condiciones de inercia e incertidumbre, de impactos de movilización, desmovilización o transferencias. En estos pares se concentra la mayor incertidumbre competitiva del sistema. A diferencia de la covarianza clásica, esta métrica es geoméricamente coherente en el simplex y no depende del nivel absoluto de cada componente.

El resumen relacional se fundamenta en la varianza de los log-ratios composicionales, calculada sobre el conjunto de escenarios proyectados. En coherencia con la geometría de Aitchison, la variación global del sistema puede resumirse mediante una medida media de las fluctuaciones relativas en el simplex.

A nivel de actor  $i$  se definen dos indicadores operativos:

$$\text{Var}_i^{\text{media}} = (1/(k - 1)) \cdot \sum_{j \neq i} \text{Var}[\log(x_i/x_j)]$$

$$\text{Var}_i^{\text{max}} = \max_{j \neq i} \text{Var}[\log(x_i/x_j)]$$

El primero aproxima su estabilidad relacional global y el segundo identifica la relación bilateral más inestable asociada al actor.

Para interpretar magnitudes sin recurrir a umbrales absolutos, se utiliza una evaluación relativa respecto a la distribución empírica transversal. Sea  $\mu_V$  y  $\sigma_V$  la media y desviación estándar de  $\text{Var}_i^{\text{media}}$ ; se define  $Z_i = (\text{Var}_i^{\text{media}} - \mu_V)/\sigma_V$ . Valores  $Z_i \leq -1$  indican anclaje estructural robusto;  $-1 < Z_i < 1$  comportamiento relacional neutro; y  $Z_i \geq 1$  sensibilidad relacional elevada. Para  $\text{Var}_i^{\text{max}}$ , se consideran tensiones bilaterales relevantes cuando el valor supera el percentil  $Q_{75}$ , y críticas cuando supera  $Q_{90}$ .

La tabla 11 resume variación relacional mediante varianzas de log-cocientes:  $T_{ij} = Var_s \left( \ln(x_i/x_j) \right)$  a través de escenarios. Valores próximos a cero indican invariancia estructural en la razón  $i/j$ ; valores elevados indican sensibilidad relativa al escenario. Esta métrica complementa CLR/ILR: no impone bloques a priori, sino que mide co-movimiento relativo por pares de actores.

**Tabla 11 — Variación relacional (Var[ln(x\_i/x\_j)])**

Partido	VarLogRatio_media	VarLogRatio_max
PP	0.021	0.050
PSOE	0.027	0.057
VOX	0.021	0.064
Adelante Andalucía	0.019	0.045
Por Andalucía	0.055	0.104
Otro	0.047	0.097
SALF	0.018	0.043
En blanco	0.040	0.104

La combinación de ambos indicadores permite distinguir configuraciones estructurales diferenciadas: estabilidad sistémica cuando  $Var_i^{media}$  es baja y  $Var_i^{max}$  no supera  $Q_{75}$ ; estabilidad global con tensiones puntuales cuando  $Var_i^{media}$  es neutra, pero  $Var_i^{max}$  elevado; y sensibilidad sistémica cuando ambos indicadores se sitúan en rangos altos. Estos indicadores capturan variación en posiciones relativas dentro del simplex y no volatilidad porcentual absoluta. En consecuencia, magnitudes elevadas deben interpretarse como movilidad relacional del equilibrio competitivo bajo incertidumbre controlada, complementando el plano aritmético del modelo mediante una medida formal de resiliencia estructural.

### Biplot CLR

El biplot CLR constituye una herramienta exploratoria que permite visualizar proximidades, oposiciones y ejes dominantes del sistema en términos de log-ratios centrados. Al estar construido sobre la transformación CLR y el PCA correspondiente, el plano PC1-PC2 representa el subespacio de máxima varianza relacional dentro de la geometría de Aitchison. Las flechas (vectores de carga) indican la dirección y magnitud de la covariación relativa de cada componente respecto al conjunto del sistema.

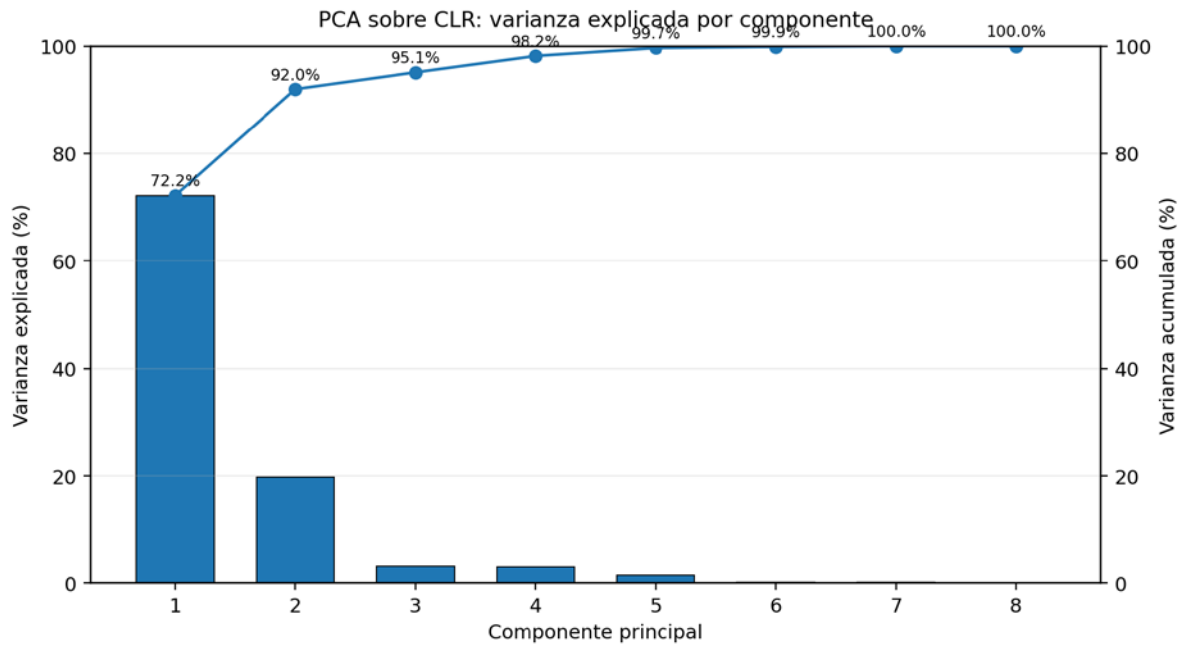
El análisis se realiza aplicando un Análisis de Componentes Principales (PCA) sobre la matriz de datos transformada en coordenadas log-ratio centradas (CLR). Sea  $X \in S^k$  el conjunto composicional de escenarios y  $Z = clr(X)$  la transformación al espacio euclídeo de dimensión  $k - 1$  inducido por la geometría de Aitchison. La matriz de covarianzas log-ratio se define como  $\Sigma_{CLR} = Cov(Z)$ , cuya descomposición espectral  $\Sigma_{CLR} = PAP^T$  permite identificar los ejes principales de variación relacional. Los autovalores  $\lambda^1 \geq \lambda^2 \geq \dots \geq \lambda_{k-1}$  contenidos en  $\Lambda$  representan la varianza explicada por cada componente principal en el subespacio CLR. En el biplot proyectado sobre el plano generado por los vectores propios asociados a  $\lambda^1$  y  $\lambda^2$ , la proporción de varianza explicada se expresa como  $(\lambda^1 + \lambda^2) / \sum_i \lambda_i$ , lo que cuantifica el grado en que la estructura relacional del sistema puede representarse adecuadamente en dos dimensiones.

Geoméricamente, las componentes principales definen un subespacio óptimo en el sentido de Aitchison, maximizando la varianza total de los log-ratios proyectados. Los puntos del biplot corresponden a escenarios en el subespacio generado por los vectores propios principales, mientras que los vectores de carga representan las direcciones de máxima covariación relativa de cada componente respecto al conjunto. En términos formales, el PCA-CLR equivale a una descomposición ortogonal del espacio clr que preserva la métrica de Aitchison, garantizando coherencia con la estructura composicional del sistema. El biplot resultante constituye, por tanto, una representación geométrica global de la variación relacional inducida por los escenarios, alineada con la teoría composicional avanzada.

En resumen, la proximidad entre observaciones sugiere similitud composicional en el espacio reducido, mientras que la separación indica configuraciones relativamente distintas en los componentes principales retenidos. La longitud y dirección de las flechas ayudan a interpretar qué componentes estructuran con mayor intensidad cada eje. Flechas más largas suelen asociarse a una mayor capacidad explicativa en la configuración del plano. La interpretación conjunta de puntos y vectores permite distinguir no solo agrupamientos de observaciones, sino también qué componentes empujan la estructura del sistema en direcciones específicas del espacio PCA.

Antes de representar conjuntamente escenarios y vectores de carga, conviene observar cuánta varianza relacional del sistema queda capturada por las primeras componentes. El scree plot permite verificar si la proyección bidimensional resume de forma suficiente la estructura composicional.

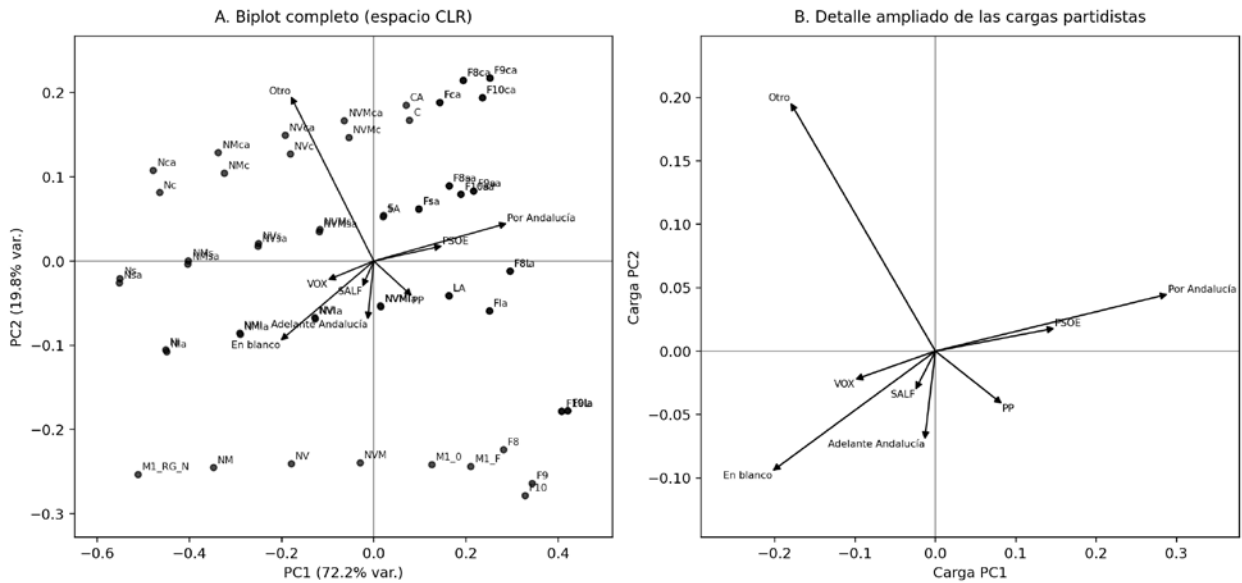
**Gráfico 2 — Varianza explicada del PCA sobre coordenadas CLR (Scree plot)**



Nota: el scree plot muestra la proporción de varianza explicada por cada componente principal del PCA aplicado a coordenadas CLR. Las barras representan la contribución individual de cada eje, mientras que la línea acumulada indica cuánta estructura relacional del sistema queda resumida a medida que se incorporan componentes sucesivos. Su lectura permite valorar si la representación en dos dimensiones ofrece una síntesis suficiente del espacio composicional.

### Gráfico 3 — Visualización exploratoria del espacio relacional (Biplot composicional: PCA sobre CLR)

PCA composicional del sistema de partidos: elecciones y cargas partidistas



Nota: el biplot muestra simultáneamente las observaciones y los vectores de carga de los componentes del análisis de componentes principales aplicado a coordenadas CLR. Los puntos resumen la posición relativa de cada observación en el espacio reducido, mientras que las flechas indican la contribución y orientación de cada componente en los dos ejes principales.  $PC1\_loading$  y  $PC2\_loading$  son las coordenadas del partido en el espacio PCA-CLR.  $x_{plot}$  e  $y_{plot}$  son las coordenadas efectivas del vector tal como se dibuja en el gráfico (escalado por los percentiles de los *scores* y por  $arrow_{scale}$ ). La columna plotted indica si el partido entra en el subconjunto que se dibuja cuando se aplica  $max_{arrows}$ .

La longitud de las flechas es proporcional a la contribución del componente a la varianza explicada en el plano proyectado. Flechas largas indican alta participación en los ejes dominantes de variación relacional; flechas cortas sugieren contribución limitada en las dos primeras componentes, aunque puedan ser relevantes en dimensiones no representadas. Por tanto, la longitud no debe interpretarse como tamaño porcentual, sino como intensidad de variación relacional capturada por el subespacio proyectado.

El ángulo entre flechas permite inferir patrones de co-movimiento en términos de log-ratios. Ángulos cercanos a  $0^\circ$  indican co-variación relativa positiva, es decir, movimiento conjunto frente al resto del sistema. Ángulos cercanos a  $180^\circ$  indican oposición relacional o redistribución relativa de masa composicional entre componentes. Ángulos próximos a  $90^\circ$  sugieren independencia relacional en el plano proyectado. Estas relaciones deben interpretarse exclusivamente dentro del marco composicional y no como correlaciones aritméticas directas.

La proyección de los puntos (escenarios) sobre la dirección de una flecha permite evaluar cómo cada escenario contribuye a la variación asociada a ese componente. Escenarios alineados con la dirección de una flecha reflejan valores relativos elevados en ese componente frente al conjunto; escenarios proyectados en dirección opuesta indican valores relativos reducidos.

La concentración de puntos alrededor del origen sugiere estabilidad estructural y bajo grado de redistribución relacional entre escenarios. Una dispersión amplia, elongaciones marcadas o agrupamientos diferenciados pueden indicar incertidumbre estructural, ejes dominantes de competencia o posibles subregímenes latentes dentro del espacio proyectado. En todo caso, el biplot CLR no establece contrastes formales, sino que sintetiza patrones geométricos globales de co-variación relativa que deben ser contrastados con los indicadores estructurales del modelo.

**Tabla 12 — Coordenadas de los partidos en el biplot CLR (PC1–PC2): loadings y vector dibujado**

Partido	partido	PC1_loading	PC2_loading	x_plot	y_plot	norm_loading	norm_plot	plotted
0	PSOE	0.3347	0.0753	0.1487	0.0181	0.3431	0.1498	1.0000
1	PP	0.1874	-0.1735	0.0832	-0.0417	0.2553	0.0931	1.0000
2	VOX	-0.2254	-0.0931	-0.1001	-0.0224	0.2439	0.1026	1.0000
3	SALF	-0.0561	-0.1272	-0.0249	-0.0306	0.1390	0.0394	1.0000
4	Adelante Andalucía	-0.0298	-0.2906	-0.0132	-0.0698	0.2921	0.0711	1.0000
5	Por Andalucía	0.6517	0.1870	0.2896	0.0449	0.6780	0.2930	1.0000
6	Otro	-0.4059	0.8156	-0.1803	0.1960	0.9110	0.2663	1.0000
7	En blanco	-0.4567	-0.3935	-0.2029	-0.0945	0.6028	0.2238	1.0000

La incorporación del plano CLR garantiza que la estimación central respete la naturaleza relativa del dato y que la estabilidad porcentual no encubra redistribuciones proporcionales internas. De este modo, la validación estructural precede a la interpretación sustantiva. A continuación, los ILR introducen contrastes dirigidos entre subconjuntos definidos teóricamente. Ambos enfoques son complementarios: CLR describe la estructura global; ILR permite contrastes confirmatorios específicos.

## II.B Análisis composicional confirmatorio (balances estructurales ILR)

### Diferenciación técnica: del rigor formal al diagnóstico estructural

La arquitectura del análisis composicional se articula en dos niveles jerárquicos complementarios, pero funcionalmente distintos. El primero, denominado 'Interpretación Formal', se sitúa en un estrato micro-analítico donde la unidad de estudio es el balance individual. Su enfoque principal es la determinación de la razón geométrica ( $g(+)/g(-)$ ), una métrica que permite cuantificar el peso relativo de un bloque sobre otro mediante coeficientes de normalización y distancias euclídeas puras. La función primordial de este nivel es la traducción: convertir el *log – ratio* abstracto en un valor multiplicativo interpretable que defina la magnitud real de la dominancia electoral.

En otro plano, el 'Diagnóstico Estructural' opera bajo una lógica macro-analítica, tomando como unidad de estudio el plano completo o la combinación ortogonal de balances. En este nivel, el interés se desplaza de la magnitud aislada hacia la identificación del régimen composicional, evaluando la trayectoria del sistema a través de los cuadrantes del simplex. Para ello, se emplean métricas sistémicas como la intensidad radial ( $r$ ) y la estabilidad direccional. Mientras que el Apartado B traduce los datos, el Apartado A asume una función evaluativa: su objetivo es detectar cambios de régimen, fenómenos de polarización relativa y tensiones compensatorias que solo emergen mediante la interacción de las coordenadas ILR.

Es fundamental distinguir entre la interpretación formal de los balances y su aplicación al diagnóstico estructural. En primer lugar, se establece la fundamentación matemática mediante coordenadas ortonormales, asegurando que cada balance sea una métrica isométrica en el espacio de Aitchison. Su propósito es la traducción sustantiva: convertir el coeficiente ILR en una razón geométrica ( $\text{Ratio\_mediana}$ ) que exprese cuántas veces es superior el bloque positivo respecto al negativo.

Esta dualidad permite una lectura integral del sistema multipartidista. Primero se garantiza el rigor matemático en la descripción de las fuerzas en presencia, para posteriormente proporcionar el marco diagnóstico necesario para discernir entre una fluctuación coyuntural y una reconfiguración estructural consolidada en el espacio de Aitchison. En resumen, primero se responde a la precisión del dato (el 'qué'), para posteriormente proporcionar la interpretación del comportamiento sistémico (el 'cómo' y el 'hacia dónde'). Ambos niveles son complementarios y necesarios para una caracterización completa de la arquitectura electoral bajo análisis.

### Definición formal de los balances ILR por bloques

En ILR se proyecta el simplex a coordenadas ortonormales mediante balances definidos por bloques (+) y (-). El programa utiliza balances ILR predefinidos para capturar contrastes sustantivos entre subconjuntos de actores (p.ej., izquierda vs derecha; tensiones intra-bloque).

Para un balance con  $r$  componentes en el bloque (+) y  $s$  en el bloque (-), el programa define:

$$b = \text{coef} \cdot \ln(g(+)/g(-)), \quad \text{coef} = \sqrt{r \cdot s / (r+s)},$$

donde  $g(+)$  y  $g(-)$  son las medias geométricas de los componentes en cada bloque. El coeficiente garantiza ortonormalidad en el espacio de Aitchison. En su interpretación:  $b > 0$  implica predominio relativo del *bloque(+)*,  $b < 0$  del *bloque(-)*, y  $b = 0$  equilibrio geométrico. La tabla 13 resume la localización central de cada balance (media, mediana, Huber) y su dispersión bajo escenarios. La coincidencia entre mediana y HuberMAD sugiere estabilidad frente a escenarios extremos; diferencias relevantes sugieren sensibilidad del contraste a condiciones específicas del espacio de escenarios (inercia/incertidumbre). Se reportan balances ILR definidos por bloques (pos vs neg). Un valor mayor indica desplazamiento relativo hacia el bloque

positivo; un valor menor, hacia el bloque negativo. Al operar en log-ratios isométricos, los balances capturan transferencia relativa entre grupos sin depender del nivel absoluto.

**Tabla 13 — Resumen robusto de balances ILR (media, mediana, Huber)**

Partido	Media	Mediana	Huber_SD	Huber_MAD
Izquierda_vs_Derecha	-0.143	-0.071	-0.132	-0.119
Izquierda_interna_PSOE_vs_Resto	1.135	1.145	1.138	1.139
Derecha_interna_PP_vs_VOX	1.660	1.671	1.668	1.671

La tabla muestra la localización central y la dispersión de cada balance ILR bajo los escenarios generados. La comparación entre media, mediana y estimadores Huber permite evaluar la estabilidad direccional del contraste entre bloques. La proximidad entre estos estimadores indica ausencia de distorsión por escenarios extremos, mientras que la dispersión asociada refleja sensibilidad del balance a las variaciones de inercia e incertidumbre. Cuando la mediana y el estimador robusto coinciden y la dispersión es reducida, el contraste entre bloques puede considerarse estructuralmente estable.

### Balances ILR — Interpretación formal (razón geométrica)

Cada balance ILR constituye una coordenada ortonormal del simplex y se define como  $b = \sqrt{rs/(r+s)} \cdot \log(g(+)/g(-))$ , donde  $r$  es el número de componentes en el bloque positivo,  $s$  el número de componentes en el bloque negativo,  $g(+)$  y  $g(-)$  las medias geométricas de los subconjuntos correspondientes, y  $\sqrt{rs/(r+s)}$  el coeficiente de normalización geométrica.

Este coeficiente no es arbitrario: garantiza ortonormalidad en el espacio de Aitchison, asegurando que cada balance tenga varianza comparable y que los ejes sean mutuamente ortogonales. Gracias a esta propiedad, las distancias euclídeas en el plano ILR equivalen a distancias de Aitchison en el simplex original. La interpretación sustantiva se realiza sobre la razón geométrica equivalente  $g(+)/g(-) = \exp(b/coef)$ . Valores positivos de  $b$  indican predominio relativo del bloque positivo; valores negativos, predominio del bloque negativo;  $b = 0$  implica equilibrio geométrico entre bloques. El balance no mide magnitudes absolutas de voto, sino relaciones proporcionales agregadas entre subconjuntos de actores.

La tabla 14 convierte cada balance en lectura sustantiva: reporta

- (i) coef y bloques (+)/(-),
- (ii)  $P(b>0)$  como estabilidad direccional del signo del contraste,
- (iii) la razón geométrica equivalente  $g(+)/g(-)=\exp(\text{mediana}/\text{coef})$ , y
- (iv) escenarios extremos (máximo/mínimo) que inducen los desplazamientos relativos más grandes.

**Tabla 14 — Balances ILR interpretables (coeficiente, direccionalidad y razón geométrica)**

Partido	Balance	Bloque (+)	Bloque (-)	coef	Mediana(b)	HuberMA D(b)	$P(b>0)$	Ratio_mediana $g(+)/g(-)$	Escenario_max	Escenario_min
0	Izquierda_vs_Derecha	PSOE, Adelante Andalucía, Por Andalucía	PP, VOX, SALF	1.225	-0.071	-0.119	0.000	0.944	F8la, F9L, F9la	Nla, Nl, M1_RG_N
1	Izquierda_interna_PSOE_vs_Resto	PSOE	Adelante Andalucía, Por Andalucía	0.816	1.145	1.139	1.000	4.063	Fs, F8a, F8sa	Nl, Nla, NMI
2	Derecha_interna_PP_vs_VOX	PP	VOX, SALF	0.816	1.671	1.671	1.000	7.739	F9la, F9L, LA	Nca, Nc, M1_RG_N

Esta tabla traduce cada coordenada ILR en términos interpretables, incorporando tres niveles de lectura: direccionalidad probabilística ( $P(b > 0)$ ), que mide la estabilidad del signo del balance; razón geométrica equivalente, que expresa el contraste en términos multiplicativos; y escenarios extremos, que identifican las combinaciones de inercia e incertidumbre responsables de los máximos desplazamientos relativos. La combinación de estos elementos permite distinguir entre predominio estructural estable y desplazamientos dependientes de condiciones específicas del espacio de escenarios.

#### **Lectura por balance**

- Izquierda\_vs\_Derecha. Balance formal:  $b = \text{coef} \cdot \log(g(+)/g(-))$ ,  $\text{coef}=1.225$ ;  $g(+)$  = media geométrica de [PSOE, Adelante Andalucía, Por Andalucía] y  $g(-)$  = media geométrica de [PP, VOX, SALF]. Mediana(b)=-0.071  $\Rightarrow$  predominio relativo del bloque (-) frente al bloque (+). Razón geométrica equivalente:  $g(+)/g(-)=\exp(\text{mediana}/\text{coef})\approx 0.944$ . Direccionalidad:  $P(b>0)=0.000$ . Extremos: b alto [F8la, F9L, F9la] | b bajo [Nla, Nl, M1\_RG\_N].
- Izquierda\_interna\_PSOE\_vs\_Resto. Balance formal:  $b = \text{coef} \cdot \log(g(+)/g(-))$ ,  $\text{coef}=0.816$ ;  $g(+)$  = media geométrica de [PSOE] y  $g(-)$  = media geométrica de [Adelante Andalucía, Por Andalucía]. Mediana(b)=1.145  $\Rightarrow$  predominio relativo del bloque (+) frente al bloque (-). Razón geométrica equivalente:

$g(+)/g(-)=\exp(\text{mediana}/\text{coef})\approx 4.063$ . Direccionalidad:  $P(b>0)=1.000$ . Extremos: b alto [Fs, F8a, F8sa] | b bajo [Nl, Nla, NMI].

- Derecha\_interna\_PP\_vs\_VOX. Balance formal:  $b = \text{coef} \cdot \log(g(+)/g(-))$ ,  $\text{coef}=0.816$ ;  $g(+)$  = media geométrica de [PP] y  $g(-)$  = media geométrica de [VOX, SALF]. Mediana(b)=1.671  $\Rightarrow$  predominio relativo del bloque (+) frente al bloque (-). Razón geométrica equivalente:  $g(+)/g(-)=\exp(\text{mediana}/\text{coef})\approx 7.739$ . Direccionalidad:  $P(b>0)=1.000$ . Extremos: b alto [F9la, F9L, LA] | b bajo [Nca, Nc, M1\_RG\_N].

## Visualización exploratoria ILR (sin criterio estructural)

La utilidad principal de este gráfico es mostrar la trayectoria del sistema en términos relacionales, no solo aritméticos. Cada desplazamiento debe leerse como una modificación en el equilibrio entre componentes de la composición. La continuidad o ruptura de la trayectoria ayuda a distinguir dinámicas graduales de reconfiguraciones más abruptas. Cuando la nube de puntos se concentra en una zona reducida, ello sugiere estabilidad relativa del sistema en los balances representados. Los hitos destacados permiten anclar la lectura comparativa entre observaciones seleccionadas. Su función no es sustituir la lectura del conjunto, sino facilitar la identificación de cambios de orientación dentro de la trayectoria global.

El plano ILR representa coordenadas ortonormales del espacio de Aitchison, esto es, una transformación isométrica del simplex composicional en un espacio euclídeo de dimensión  $D - 1$ . Cada eje corresponde a un balance formal definido como  $b = \sqrt{\left(\frac{r \cdots}{(r + s)}\right) \cdot \ln(g(+)/g(-))}$ , donde  $g(+)$  y  $g(-)$  son las medias geométricas de los subconjuntos de actores definidos en el numerador y denominador del balance, y  $r$  y  $s$  el número de componentes en cada subconjunto. El coeficiente garantiza ortonormalidad y preservación de distancias en la métrica de Aitchison.

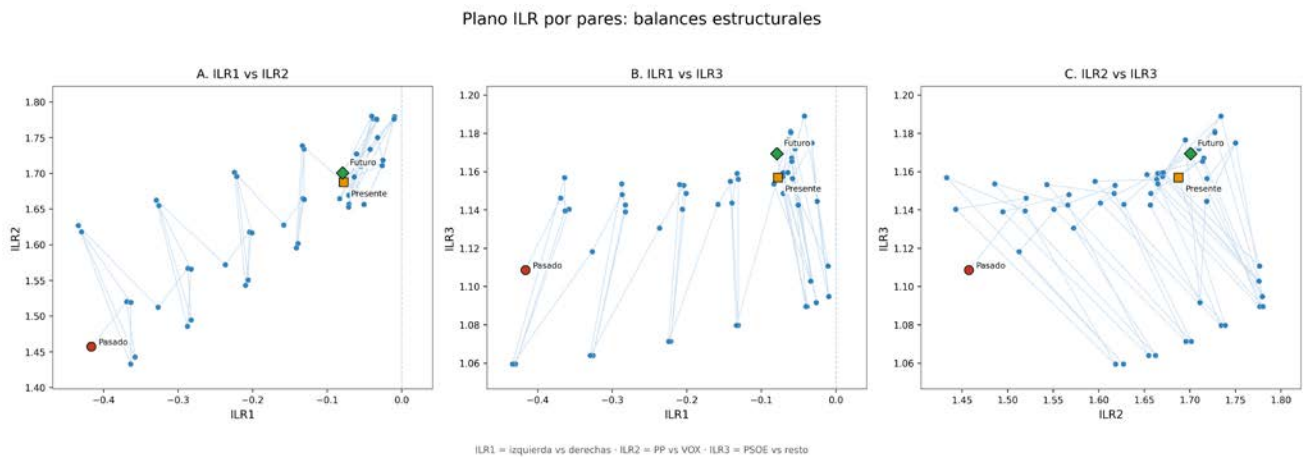


Gráfico 4 — Proyección de Balances

Nota: la figura proyecta las observaciones en tres planos bivariados construidos a partir de balances ILR. Cada punto representa una composición y las líneas muestran la trayectoria entre observaciones sucesivas. La posición relativa en cada plano resume tensiones entre bloques composicionales y permite identificar desplazamientos, estabilidad o cambios de dirección. Un valor  $b > 0$  indica predominio relativo multiplicativo del *bloque(+)* respecto al *bloque(-)*;  $b < 0$  indica el predominio inverso. El valor absoluto  $|b|$  mide la intensidad logarítmica del desequilibrio relativo. La interpretación es siempre multiplicativa (razones geométricas) y no aditiva porcentual. Las líneas en 0 delimitan regiones de signo que permiten interpretar combinaciones de dominancia relativa. Los cuadrantes describen dirección (signo) del balance, no magnitud del desplazamiento.

Panel (Izquierda\_vs\_Derecha vs PP\_vs\_VOX): permite observar simultáneamente el balance ideológico agregado y la redistribución interna dentro del bloque derecho, separando desplazamientos globales de compensaciones intra-bloque.

Panel (Izquierda\_vs\_Derecha vs PSOE\_vs\_Resto): distingue escenarios donde el predominio del bloque de izquierda se acompaña -o no- de centralidad relativa del PSOE dentro de dicho bloque.

Panel (PP\_vs\_VOX vs PSOE\_vs\_Resto): aísla tensiones internas bilaterales relativamente independientes del eje ideológico global.

Lectura por cuadrantes:  $(X>0, Y>0)$  indica predominio relativo del bloque (+) en ambos balances; el resto de los cuadrantes se interpretan análogamente cambiando el signo.

### Diagnóstico estructural ILR (cuadrantes e intensidad)

Este diagnóstico evalúa la estabilidad del sistema mediante dos dimensiones: la ocupación de cuadrantes (signo de los balances) y la intensidad del desplazamiento radial  $r = \sqrt{x^2 + y^2}$ . Mientras que los cuadrantes identifican el régimen de equilibrio relativo entre bloques, la magnitud  $r$  mide la distancia respecto al centro de equilibrio geométrico. Esta aproximación permite distinguir entre estabilidad direccional, tensión interna y reconfiguración estructural del sistema composicional. El diagnóstico ILR está definido sobre balances (contrastes por bloques), por lo que no mide niveles absolutos de voto sino transferencias relativas agregadas entre subconjuntos. En consecuencia, un cambio de régimen ILR describe reconfiguración relacional (quién gana relativamente dentro del sistema), no necesariamente un cambio grande en puntos porcentuales.

**Tabla 15 — Diagnóstico ILR por panel: probabilidad por cuadrante e intensidad (r)**

Partido	Panel	P(QI)	P(QII)	P(QIII)	P(QIV)	Mediana_r	MAD_r	P90_r	Max_r
0	Izquierda_vs_Derecha × Derecha_interna_PP_vs_VOX	0.000	1.000	0.000	0.000	1.687	0.048	1.776	1.781
1	Izquierda_vs_Derecha × Izquierda_interna_PSOE_vs_Resto	0.000	1.000	0.000	0.000	1.161	0.017	1.188	1.213
2	Derecha_interna_PP_vs_VOX × Izquierda_interna_PSOE_vs_Resto	1.000	0.000	0.000	0.000	2.030	0.056	2.093	2.108

Para cada panel (balance X × balance Y) se calculan probabilidades por cuadrante:  $QI(X > 0, Y > 0)$ ,  $QII(X < 0, Y > 0)$ ,  $QIII(X < 0, Y < 0)$ ,  $QIV(X > 0, Y < 0)$  incluyendo la  $Mediana_r$ ,  $MAD_r$ ,  $P90_r$  (percentil90) y  $Max_r$  a través de escenarios. La interpretación de los resultados de la Tabla 15 debe regirse por los siguientes criterios: una concentración elevada de probabilidad en un único cuadrante (P approx 1.000) indica estabilidad direccional plena y persistencia del régimen de signos. Por el contrario, la dispersión de probabilidad entre cuadrantes opuestos sugiere una inversión de los balances y, por tanto, una posible transición estructural. Respecto a la intensidad, una mediana\_r elevada junto a una MAD\_r reducida señala un desplazamiento consistente hacia un nuevo estado de equilibrio, mientras que valores altos de  $P90_r$  (percentil90) o  $Max_r$  advierten sobre la existencia de escenarios con comportamientos extremos o volátiles.

En síntesis, incrementos en la intensidad radial sin cambio de cuadrante representan episodios de tensión compensatoria donde el sistema refuerza su posición actual. Solo cuando el aumento de intensidad coincide con una migración de la masa probabilística hacia cuadrantes alternativos se puede hablar de una reconfiguración de la estructura relacional del modelo analizado.

### Plano ILR por pares (diagnóstico estructural)

El bloque estructural ILR incorpora criterios diagnósticos adicionales a la visualización exploratoria. Dado que la representación ILR es isométrica en el espacio de Aitchison, el gráfico ILR se estructura en torno al plano estructural con delimitación formal de cuadrantes, centroide robusto y escenarios extremos; el plano estructural integra toda la información geométrica exploratoria y añade criterios diagnósticos explícitos: ocupación por cuadrantes, estabilidad direccional e intensidad radial. En consecuencia, la lectura gráfica se orienta directamente a la detección de regímenes composicionales y tensiones compensatorias.

Criterio de estabilidad direccional: existe estabilidad cuando la mayoría sustantiva de escenarios se concentra en uno o dos cuadrantes adyacentes, no existe ocupación relevante en cuadrantes opuestos y el centroide robusto permanece en la misma región de signo. En tal caso no se observa inversión del balance composicional. Si la masa de escenarios se concentra en uno o dos cuadrantes adyacentes y no hay ocupación relevante de cuadrantes opuestos, no se observa cambio de régimen composicional.

Criterio geométrico de estabilidad: la presencia sostenida en un único cuadrante indica coherencia direccional del balance. La ocupación relevante en cuadrantes opuestos (por ejemplo QI y QIII) sugiere inversión de signo y posible transición composicional. La intensidad radial  $r = \sqrt{(b_1^2 + b_2^2)}$  mide la magnitud del alejamiento respecto al equilibrio ILR; incrementos de  $r$  sin cambio de cuadrante indican ampliación del desequilibrio interno (tensión compensatoria), mientras que cambios de cuadrante acompañados de  $r$  elevado apuntan a reconfiguración estructural.

Criterio estructural de confirmación: puede hablarse de cambio de régimen cuando la ocupación en cuadrantes opuestos se acompaña de desplazamiento significativo del centroide robusto hacia una región de signo contrario y evidencia de multimodalidad (según Dip y coeficiente de bimodalidad). Este contraste permite diferenciar fluctuación geométrica transitoria de transición estructural consolidada. Criterio de tensión compensatoria: si aumenta significativamente la intensidad radial  $r$  pero no cambia el signo dominante del balance, el sistema experimenta polarización relativa sin transición estructural.

Este gráfico está orientado al diagnóstico estructural más que a la secuencia temporal. Su propósito es mostrar cómo se distribuyen las observaciones alrededor de un centro robusto y qué grado de concentración o dispersión presenta el sistema. La posición de las observaciones respecto a las medianas permite interpretar la dirección dominante de los balances. La acumulación en determinados cuadrantes puede sugerir patrones persistentes de desequilibrio entre componentes. La comparación entre paneles permite identificar qué

balances discriminan mejor las configuraciones observadas. Un panel más disperso suele indicar mayor capacidad para separar escenarios o mayor variabilidad en esa dimensión concreta.

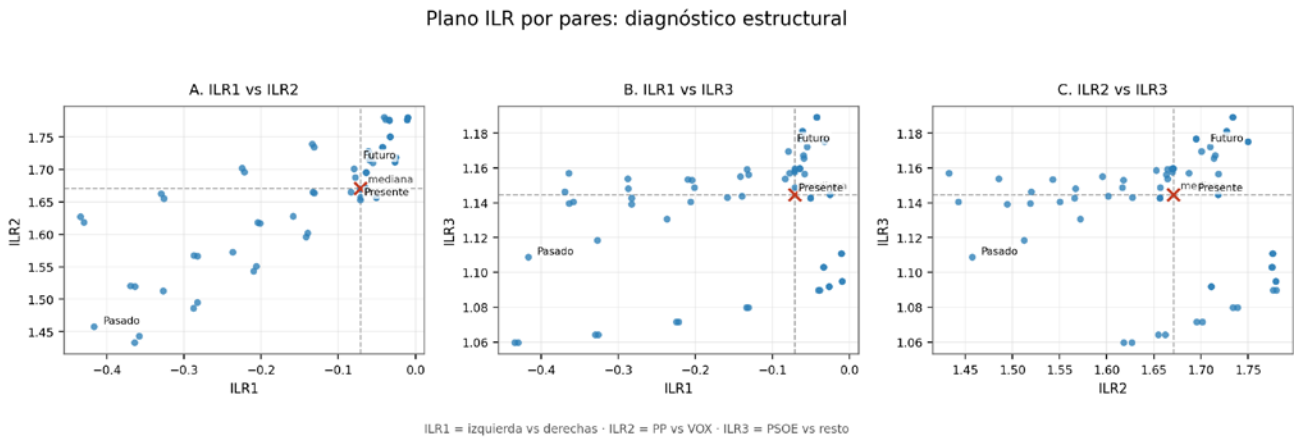


Gráfico 5 — Triángulo ILR estructural: distribución por cuadrantes,

Nota: la figura representa la distribución de observaciones en los tres planos formados por los balances ILR principales. Las líneas de mediana dividen cada panel en cuadrantes y la marca central resume el centro robusto de la distribución. Esta representación facilita la lectura de dispersión, asimetría y posición relativa del conjunto.

Cada eje representa un balance ILR definido como  $b = coef \cdot \log(g(+)/g(-))$ , donde  $g(+)$  y  $g(-)$  son las medias geométricas de los bloques comparados. Un valor  $b > 0$  indica predominio relativo del *bloque(+)* respecto al *bloque(-)*, mientras que  $b < 0$  señala predominio del *bloque(-)*. Los ejes dividen el plano en cuatro cuadrantes definidos exclusivamente por los signos de los balances, lo que permite identificar regímenes direccionales estables (cuando la ocupación se concentra en un único cuadrante) o posibles inversiones de régimen (cuando existe presencia relevante en cuadrantes opuestos). El sombreado no implica intensidad sino orientación relativa.

La magnitud del desplazamiento se interpreta mediante la distancia radial  $r = \sqrt{(b_1^2 + b_2^2)}$ , que mide el alejamiento respecto al equilibrio geométrico (origen ILR). El marcador ' $x'$ ' representa el centroide robusto (mediana), entendido como la posición estructural típica bajo el conjunto de escenarios, no como promedio aritmético. Las etiquetas identifican los escenarios con mayor distancia radial, es decir, aquellos que intensifican el desequilibrio composicional sin que ello implique necesariamente cambio de signo en los balances.



bloques de izquierda y derecha, con una influencia marginal pero constante del polo nacionalista en términos de volumen electoral agregado. La trayectoria que une los puntos 'Pasado', 'Presente' y 'Futuro' traza el vector de desplazamiento del sistema: un movimiento que permite diagnosticar si existe una tendencia de polarización hacia un vértice o una convergencia hacia el baricentro del sistema (punto central de equilibrio).

Este enfoque ternario facilita la comprensión de la estabilidad del sistema sin la abstracción de los balances log-ratio, ofreciendo una lectura intuitiva de la dominancia electoral. No obstante, su validez descansa en el proceso de *re – cierre*, que garantiza que las proporciones sumen 1 (100%) dentro de esta nueva subcomposición, permitiendo así comparar la fuerza relativa de los bloques prescindiendo de otras variables menores o ruido estadístico.

### Síntesis interpretativa de escenarios: variaciones y contribuciones estructurales

A partir de los tres escenarios de referencia -Pasado, Presente y Futuro-, se incorporan dos visualizaciones complementarias orientadas a sintetizar el sentido sustantivo del desplazamiento composicional del sistema. La primera representa la variación de cada partido entre escenarios en puntos porcentuales, permitiendo identificar ganancias, pérdidas y estabilidad relativa de forma inmediata. La segunda descompone la distancia de Aitchison entre pares de escenarios para estimar qué componentes explican en mayor medida el cambio geométrico total del sistema.

Este gráfico está pensado como una síntesis descriptiva de cambios entre configuraciones seleccionadas. Su principal ventaja es condensar mucha información en una matriz de lectura rápida, útil para identificar patrones de variación de forma inmediata. La comparación entre columnas permite distinguir cambios ya producidos, cambios esperados y cambio acumulado entre escenarios extremos. De ese modo, el heatmap separa movimientos de corto alcance de transformaciones más consistentes. Su interpretación se complementa con los análisis composicionales estructurales, ya que muestra diferencias directas entre escenarios, pero no por sí solo la relevancia geométrica de cada cambio dentro del sistema completo.

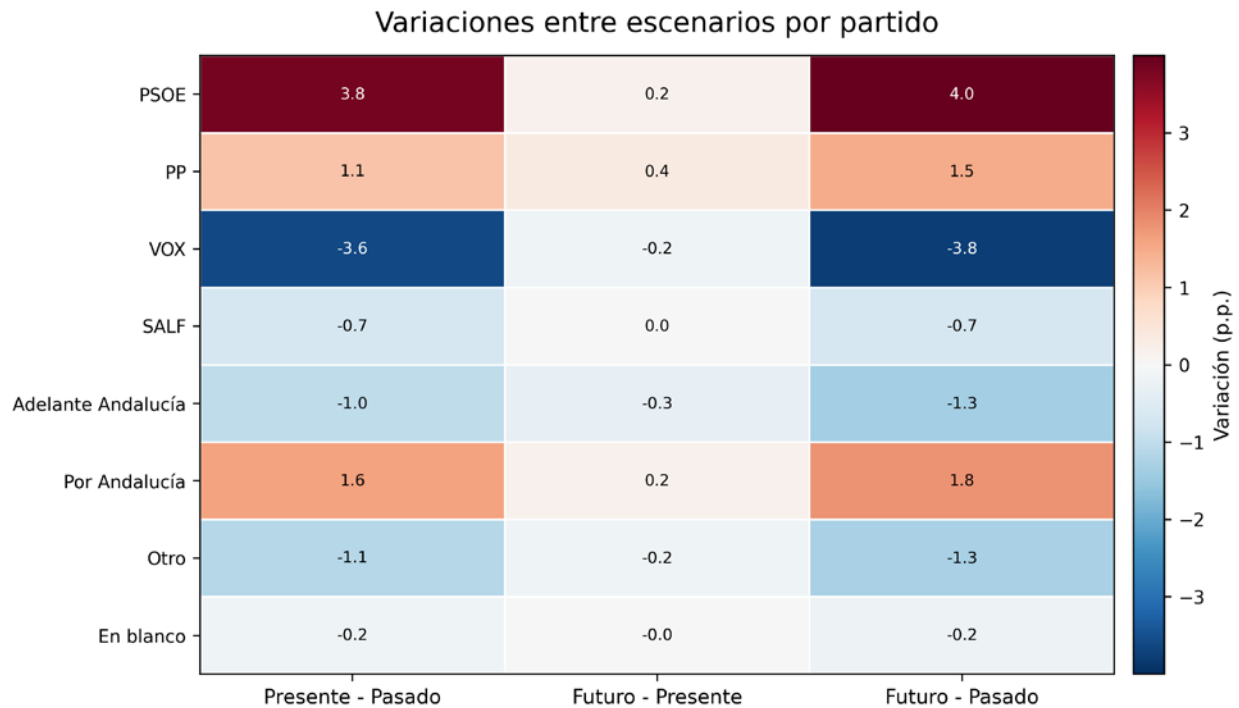


Gráfico 7 - Variaciones entre escenarios por partido

El heatmap resume las variaciones entre escenarios para cada componente de la composición. Las filas corresponden a componentes o partidos y las columnas a comparaciones entre escenarios. Los colores indican dirección e intensidad del cambio y permiten identificar de forma inmediata ganancias, pérdidas y estabilidad relativa.

La comparación Presente - Pasado resume el desplazamiento ya producido respecto al escenario de referencia anterior.

La comparación Futuro - Presente representa la dirección esperada del ajuste prospectivo del sistema.

La comparación Futuro - Pasado sintetiza el cambio acumulado y permite distinguir reconfiguraciones graduales de desplazamientos estructurales más nítidos.

El heatmap permite identificar de forma inmediata qué partidos concentran los movimientos sustantivos entre escenarios. Su utilidad principal reside en traducir el espacio composicional a diferencias legibles en términos políticos, sin perder la referencia comparativa entre bloques y actores territoriales. A diferencia de los planos ILR o del biplot CLR, esta figura no representa relaciones geométricas entre componentes, sino cambios directos entre configuraciones seleccionadas del sistema.

La principal utilidad del gráfico 6 es distinguir entre cambios visibles y cambios estructuralmente decisivos. No todos los movimientos tienen el mismo peso en la geometría del sistema, y esta descomposición permite identificar qué componentes explican realmente la distancia total. La comparación entre paneles permite observar si los motores del cambio se mantienen estables o si varían según la transición analizada. Ello resulta especialmente útil para separar transformaciones acumulativas de ajustes más específicos de una comparación concreta. Este tipo de figura complementa bien los gráficos descriptivos de variación, porque traduce

diferencias entre escenarios en términos de relevancia composicional. Así, ayuda a interpretar qué componentes alteran de forma más profunda la estructura relacional del conjunto.

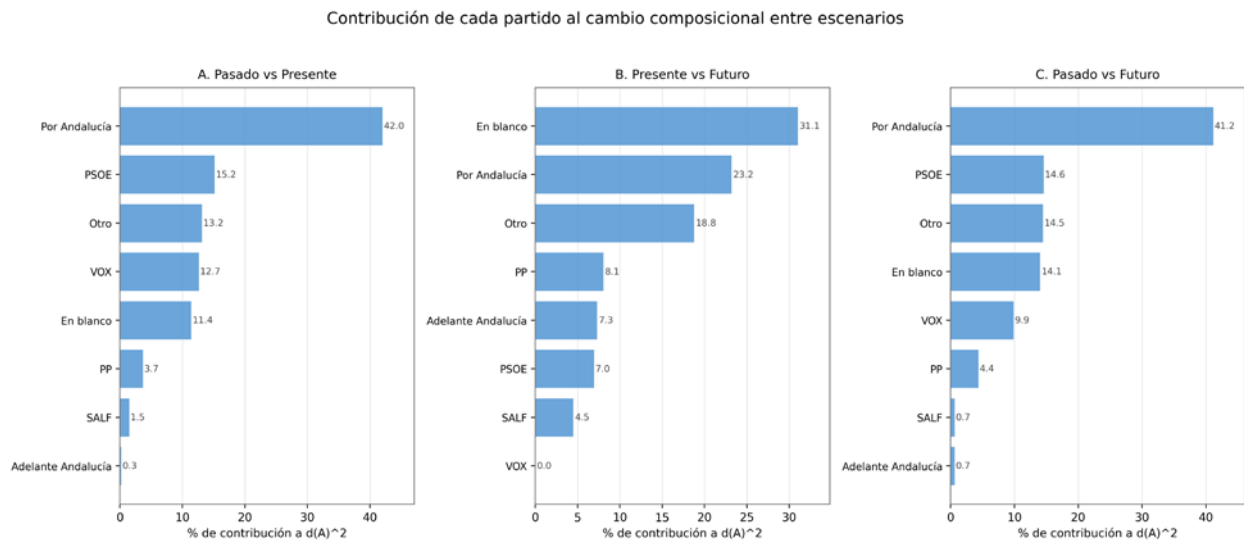


Gráfico 8 - Contribución de cada partido al cambio composicional entre escenarios

Las barras expresan el porcentaje de contribución de cada partido a la distancia de Aitchison al cuadrado entre dos escenarios. Una mayor contribución indica que la variación relativa de ese componente explica una fracción más alta del cambio geométrico total. La figura descompone la distancia de Aitchison al cuadrado entre pares de escenarios en la contribución porcentual de cada componente. Una barra más alta indica que ese componente explica una fracción mayor del desplazamiento composicional total entre las configuraciones comparadas.

El panel Pasado vs Presente identifica qué partidos explican la reconfiguración ya observada del sistema.

El panel Presente vs Futuro señala qué actores sostienen la proyección estructural del escenario prospectivo.

El panel Pasado vs Futuro resume la transformación composicional total acumulada entre ambos extremos.

La descomposición de la distancia composicional aporta una lectura más estructural que el heatmap. No se limita a indicar cuánto cambia cada partido en términos aritméticos, sino que estima qué componentes son responsables de la mayor parte del desplazamiento relacional del sistema en la geometría de Aitchison. De este modo, permite distinguir entre variaciones visibles pero periféricas y variaciones que alteran efectivamente la posición composicional conjunta del sistema de partidos.

La lectura combinada de ambas figuras permite separar dos niveles analíticos. El heatmap resume la dirección y magnitud de los cambios entre escenarios en el plano político inmediato; la descomposición de Aitchison identifica qué partidos convierten esas variaciones en un cambio geométrico sistémico relevante. En conjunto, ambos gráficos funcionan como síntesis interpretativa del bloque de escenarios antes de pasar a la validación bootstrap de los estimadores centrales.

## Validación bootstrap de los estimadores centrales bajo un esquema de re-muestreo de escenarios completos

La función de esta sección es añadir una capa de validación inferencial al análisis descriptivo y estructural. No se limita a mostrar una posición central, sino que examina su estabilidad ante múltiples réplicas del conjunto de observaciones. Una baja dispersión bootstrap sugiere que la estimación central es robusta y poco dependiente de casos concretos. Una dispersión alta, en cambio, invita a interpretar la configuración con mayor cautela. La validación bootstrap es especialmente útil cuando se desea distinguir entre patrones consistentes del sistema y resultados que podrían depender en exceso de una estructura observacional concreta.

Se presenta la validación bootstrap de los estimadores centrales bajo un esquema de re-muestreo de escenarios completos. A diferencia de un bootstrap clásico sobre observaciones individuales, el procedimiento replica filas íntegras del sistema estimado, preservando la estructura interna de cada escenario y, con ello, la interdependencia entre partidos y la restricción composicional de suma 100. Este enfoque garantiza que la incertidumbre evaluada sea estructural y no meramente aritmética. Para cada partido se reporta la media de las réplicas bootstrap y, entre paréntesis, el intervalo de confianza empírico (percentiles 5%-95%). La comparación entre media, mediana y estimadores robustos (Huber con escala SD y Huber con escala MAD) permite diagnosticar la estabilidad distributiva del centro bajo variabilidad de escenarios.

La proximidad entre los cuatro estimadores en la mayoría de los partidos indica distribuciones bootstrap aproximadamente simétricas y unimodales, sin valores extremos influyentes ni colas asimétricas relevantes. Cuando media, mediana y estimadores robustos convergen y los intervalos de confianza son estrechos, puede interpretarse que el sistema se encuentra en un régimen probabilístico estable, donde pequeñas perturbaciones en la selección de escenarios no alteran sustancialmente la posición central estimada. Asimismo, la anchura de los intervalos refleja la sensibilidad estructural del partido a la variabilidad entre escenarios completos. Partidos con intervalos relativamente más amplios presentan mayor elasticidad composicional, es decir, mayor dependencia de la configuración global del sistema.

En cambio, intervalos estrechos indican alta inercia estructural y consistencia bajo re-muestreo. En conjunto, la tabla permite evaluar simultáneamente tres dimensiones:

- (i) coherencia entre estimadores centrales,
- (ii) estabilidad frente a perturbaciones estructurales y
- (iii) magnitud de la incertidumbre composicional.

Esta validación no sólo cuantifica precisión estadística, sino que confirma si la estimación recomendada se mantiene robusta cuando se preserva la arquitectura geométrica completa del sistema electoral.

**Tabla 16a — Validación bootstrap en espacio composicional (inercia e incertidumbre)**

Partido / Escenario	Media	Mediana	Huber_SD	Huber_MAD
PP	44.09 (43.79-44.40)	43.83 (43.65-44.18)	44.04 (43.71-44.38)	44.01 (43.66-44.38)
PSOE	25.59 (25.26-25.92)	26.05 (25.71-26.34)	25.74 (25.40-26.06)	25.81 (25.41-26.22)
VOX	9.85 (9.61-10.10)	9.34 (9.14-9.83)	9.75 (9.50-10.02)	9.64 (9.31-10.01)
Adelante Andalucía	7.05 (6.95-7.15)	6.96 (6.89-7.07)	7.02 (6.92-7.13)	7.01 (6.91-7.12)
Por Andalucía	5.81 (5.64-5.98)	5.92 (5.69-6.07)	5.86 (5.67-6.04)	5.84 (5.65-6.02)
Otro	3.78 (3.63-3.93)	3.81 (3.66-4.00)	3.79 (3.62-3.95)	3.79 (3.62-3.95)
SALF	3.41 (3.36-3.47)	3.37 (3.34-3.40)	3.39 (3.34-3.45)	3.39 (3.34-3.45)
En blanco	0.42 (0.40-0.43)	0.40 (0.37-0.43)	0.41 (0.39-0.43)	0.41 (0.39-0.43)

Nota: la validación bootstrap evalúa la estabilidad de los estimadores centrales bajo re-muestreo repetido de escenarios completos. Los resultados permiten valorar la robustez de las posiciones estimadas y la sensibilidad del análisis frente a variaciones de la muestra observada.

La tabla 16b resume la validación bootstrap desde la perspectiva de la dispersión absoluta (*SD*) y la volatilidad relativa (*CV*) en el espacio composicional. A diferencia de la tabla anterior, centrada en la posición del estimador, aquí se evalúa la magnitud de la variabilidad estructural inducida por el re-muestreo de escenarios completos. Los valores reportados corresponden a la media bootstrap y, entre paréntesis, al intervalo empírico 5, lo que permite cuantificar no solo la dispersión media sino también su estabilidad bajo perturbaciones del sistema. La desviación estándar (*SD*) mide la amplitud absoluta de la variación porcentual bajo re-muestreo estructural.

Valores más elevados indican mayor sensibilidad del partido a la configuración global de escenarios. Sin embargo, dado que la magnitud electoral de los partidos es heterogénea, la comparación directa de *SD* puede inducir interpretaciones sesgadas. Por ello se incorpora el coeficiente de variación (*CV*), definido como  $SD/mediana$ , que expresa la volatilidad relativa respecto al tamaño del partido. El *CV* permite distinguir entre inercia estructural y dependencia contextual. Valores inferiores al 10% indican que la posición estimada del partido es altamente consistente bajo distintas combinaciones de escenarios completos, reflejando consolidación estructural. En cambio, valores superiores sugieren que la estimación es más elástica y depende en mayor medida del equilibrio composicional entre bloques.

**Tabla 16b — Validación bootstrap en espacio composicional (inercia e incertidumbre)**

Partido / Escenario	SD	CV
PP	1.47 (1.28-1.66)	3.3% (2.9-3.8)
PSOE	1.58 (1.30-1.85)	6.2% (5.0-7.3)
VOX	1.15 (1.00-1.30)	11.7% (10.3-13.0)
Adelante Andalucía	0.48 (0.41-0.55)	6.8% (5.9-7.8)
Por Andalucía	0.81 (0.72-0.90)	14.0% (12.2-15.8)
Otro	0.72 (0.63-0.82)	19.2% (16.5-21.8)
SALF	0.27 (0.23-0.31)	7.8% (6.7-8.9)
En blanco	0.08 (0.07-0.09)	18.5% (16.7-20.3)

En los resultados observados, los partidos que presentan SD moderadas y CV relativamente bajos, expresan estabilidad bajo perturbaciones del sistema. Formaciones intermedias que muestran coeficientes de variación contenidos son compatibles con una posición estable, aunque de menor magnitud absoluta. Por el contrario, partidos de menor tamaño o con base territorial específica exhiben CV elevados. En estos casos, la volatilidad relativa no necesariamente implica inestabilidad política, sino sensibilidad composicional: pequeñas variaciones en el equilibrio global del sistema generan cambios proporcionalmente más intensos en su estimación.

Particularmente relevante es el caso de partidos con CV superiores al 20%, donde la posición estimada depende fuertemente de la configuración estructural del resto de actores. Esta situación sugiere baja inercia composicional y alta dependencia del entorno sistémico, lo que exige cautela interpretativa en escenarios de alta incertidumbre. En conjunto, la lectura combinada de SD y CV permite separar tres perfiles:

- (i) coherencia entre estimadores centrales,
- (ii) actores estables pero sensibles a la arquitectura global del sistema, y
- (iii) actores altamente dependientes del equilibrio composicional, cuya estimación es más susceptible a reconfiguraciones de régimen.

Esta diferenciación refuerza el diagnóstico de inercia e incertidumbre desde una perspectiva geométrica coherente con la restricción de suma 100 y la interdependencia estructural del espacio electoral.

## II.C. Diagnóstico sistémico ILR ortonormal (base completa)

El diagnóstico en espacio composicional se realiza mediante una doble aproximación:

- (i) análisis exploratorio en coordenadas CLR y
- (ii) contraste confirmatorio en base ILR ortonormal completa.

Esta estrategia permite evaluar la estabilidad relativa del sistema respetando la restricción de suma constante y la interdependencia estructural entre partidos. En coordenadas CLR, cada componente se define como  $CLR_i(x) = \log(x_i/g(x))$ , donde  $g(x)$  es la media geométrica del vector completo. Esta transformación expresa la posición relativa de cada partido respecto al centro geométrico del sistema. Valores positivos indican que el partido se sitúa sistemáticamente por encima de la media geométrica global (ventaja relativa estructural), mientras que valores negativos reflejan posición estructuralmente inferior en términos proporcionales. A diferencia del porcentaje bruto, el CLR captura la relación interna entre todos los componentes simultáneamente.

La tabla presenta, para cada partido, la media CLR y el estimador robusto CLR\_HuberMAD, junto con métricas de forma distributiva (Dip test de unimodalidad, coeficiente de bimodalidad BC), medida de asimetría estructural (M) y estabilidad relativa penalizada ( $IRP^*$ ). La coincidencia entre CLR\_mean y CLR\_Huber MAD indica ausencia de colas extremas o valores influyentes en el espacio log-ratio. Valores elevados de  $IRP^*$  CLR, combinados con M CLR cercano a cero, señalan estabilidad relacional robusta bajo re-muestreo de escenarios completos. Desde el punto de vista interpretativo, un partido con CLR alto y estable no solo presenta nivel porcentual elevado, sino predominio estructural consistente en el equilibrio proporcional del sistema. Por el contrario, inestabilidad CLR -manifestada en multimodalidad, asimetría significativa o bajo  $IRP^*$ - sugiere dependencia de reequilibrios del vector completo, es decir, sensibilidad a cambios en la arquitectura interna del sistema más allá de simples oscilaciones aritméticas.

Los resultados observados muestran predominio relacional estable en los partidos principales con distribuciones unimodales, asimetría prácticamente nula y altos valores de  $IRP^*$ . Esto indica que su posición proporcional se mantiene consistente bajo perturbaciones estructurales. En contraste, algunos partidos territoriales o de menor magnitud presentan señales de multimodalidad o mayor asimetría lo que sugiere mayor sensibilidad a configuraciones alternativas del sistema composicional.

Tabla 17 — CLR por partido (log-ratio respecto a la media geométrica)

Partido	CLR_mean	CLR_HuberMAD	CLR_Dip_p	CLR_Modalidad	CLR_BC	CLR_M	CLR_IRP*
PP	1.955	1.960	0.590	Unimodal	0.544	0.000	0.966
PSOE	1.410	1.418	0.676	Unimodal	0.657	0.230	0.718
VOX	0.450	0.440	0.389	Unimodal	0.631	0.171	0.721
Adelante Andalucía	0.120	0.124	0.040	Multimodal	0.661	0.239	0.523
Por Andalucía	-0.081	-0.069	0.846	Unimodal	0.619	0.145	0.216
Otro	-0.521	-0.499	0.668	Unimodal	0.584	0.065	0.722
SALF	-0.607	-0.609	0.273	Unimodal	0.504	0.000	0.923
En blanco	-2.725	-2.727	0.982	Unimodal	0.554	0.000	0.939

La combinación de análisis exploratorio CLR y balances ILR confirmatorios permite distinguir entre estabilidad geométrica global del sistema y tensiones estructurales específicas entre bloques. Mientras el CLR evalúa la posición relacional individual respecto al centro geométrico, los balances ILR analizan comparaciones ortonormales entre subconjuntos de actores, proporcionando una lectura estructural por ejes independientes. Es fundamental subrayar que las métricas CLR e ILR operan exclusivamente como diagnóstico relacional. No intervienen en la selección automática del estimador óptimo, que se determina en el plano aritmético mediante el criterio *Eff* \*. Su función es contextualizar geoméricamente la estimación porcentual, asegurando que la estabilidad observada en niveles absolutos no oculte reequilibrios proporcionales internos del sistema. De este modo, se evita confundir estabilidad aritmética con estabilidad composicional real y se garantiza coherencia entre magnitud porcentual y estructura geométrica subyacente.

## Reglas automáticas composicionales (CLR/ILR): validación estructural del estimado

Las reglas composicionales no intervienen en la selección del estimador central. Su finalidad es detectar posibles inconsistencias entre estabilidad aritmética y estabilidad relacional. Son diagnósticas y no alteran la selección del centro en el bloque aritmético. Operan como validación estructural posterior bajo clausura ( $\Sigma=100$ ), distinguiendo estabilidad porcentual (plano aritmético) de estabilidad relacional (plano log-ratio).

Criterio A — Consistencia estructural:

si IRP\* es elevado y simultáneamente CLR\_IRP\* es elevado y CLR\_M es bajo, la estabilidad porcentual se considera estructuralmente coherente

Criterio B — Tensión compensatoria:

si IRP\* es elevado, pero se observa incremento de intensidad radial ILR ( $r$ ) sin cambio de cuadrante, la variación se interpreta como redistribución interna compensatoria y no como cambio de régimen.

Criterio C — Posible transición estructural:

si existe ocupación relevante de cuadrantes opuestos en el plano ILR y simultáneamente CLR\_M se incrementa, puede existir reconfiguración relacional del sistema. Esta señal no invalida el centro aritmético, pero exige cautela interpretativa.

Reglas automáticas de estabilidad y detección de cambio de régimen mediante extensión composicional CLR/ILR. Umbrales usados por el programa: IRP\* estable si  $\geq 0.75$ ; M estable si  $\leq 0.20$ ; inestabilidad relativa CLR si  $CLR\_IRP^* < 0.45$  o si  $CLR\_M > 0.40$ ; balance ILR crítico por intensidad si  $ILR\_M \geq 0.60$ .

Criterio operativo: IRP\* alto (estabilidad porcentual) y M bajo (baja fragmentación) pero con inestabilidad relacional en CLR (CLR\_IRP\* bajo y/o CLR\_M alto). Interpretación: la estabilidad en porcentaje puede estar sostenida por compensaciones internas en el vector composicional (cambios en relaciones relativas), sin desplazamiento neto visible en %.

### Regla 1 — Estabilidad en % con reequilibrio composicional (CLR inestable)

Se detectan actores con estabilidad estructural en % (IRP\* alto, M bajo, unimodalidad) pero inestabilidad relativa en CLR. Esto sugiere que la estabilidad porcentual puede estar sostenida por compensaciones en el vector composicional (cambios en relaciones relativas).

- Adelante Andalucía
- Por Andalucía

#### Detalle de la Regla 1 (estabilidad % vs inestabilidad CLR)

Partido	IRP_star	M	CLR_IRP*	CLR_M	CLR_Dip_p
Adelante Andalucía	0.945	0.000	0.523	0.239	0.040
Por Andalucía	0.819	0.036	0.216	0.145	0.846

Nota:  $IRP^* \in [0,1]$  valores próximos a 1 indican alta robustez del estimador.  $M \in [0,1]$  refleja intensidad estructural (valores próximos a 0 indican segmentación; valores altos mayor fragmentación). En el plano composicional, CLR\_IRP\* mide estabilidad relativa en log-ratios, CLR\_M la intensidad estructural relacional, y CLR\_Dip\_p corresponde al p-valor de la prueba de unimodalidad ( $p < 0,05$  indica evidencia de multimodalidad).

IRP\* mide resiliencia del estimador central frente a variabilidad de escenarios en el espacio porcentual. CLR\_IRP\* mide estabilidad de la posición relativa del actor respecto a la media geométrica del sistema. Los balances ILR evalúan contrastes dirigidos entre subconjuntos de actores. Cada indicador describe una dimensión distinta del sistema.

La arquitectura multinivel del modelo garantiza que la selección del estimador óptimo se realice exclusivamente en el plano aritmético, mientras que el plano composicional opera como validación estructural posterior.

De este modo, el modelo preserva separabilidad analítica entre estimación central y diagnóstico estructural, evitando que consideraciones geométricas alteren la decisión óptima basada en eficiencia y robustez aritmética.

### III. Gráficos Distribuciones bootstrap de estimadores de localización central

Gráficos ( $B = 5000$ , IC 95%):

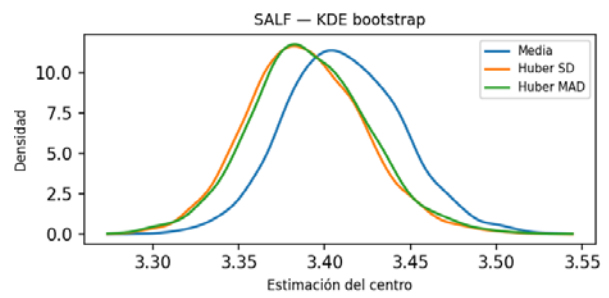
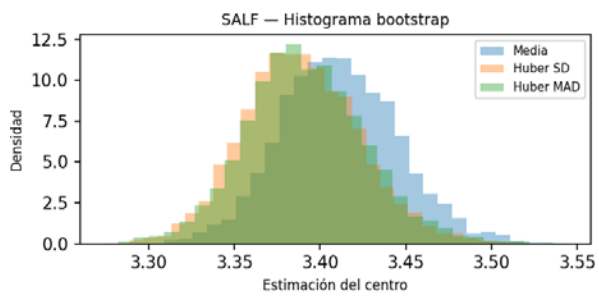
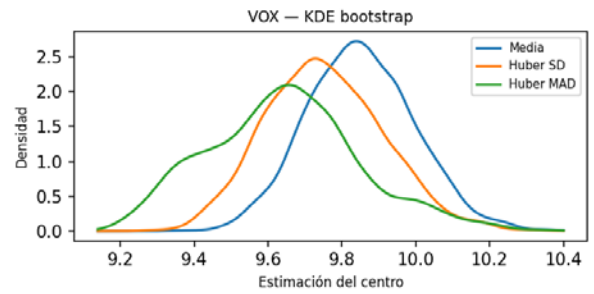
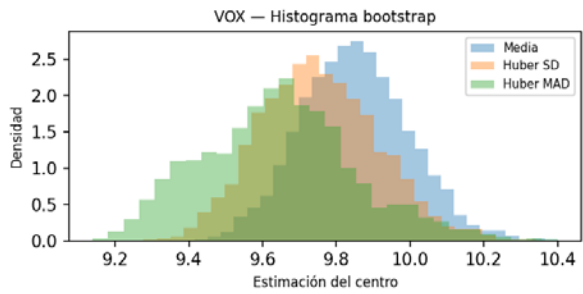
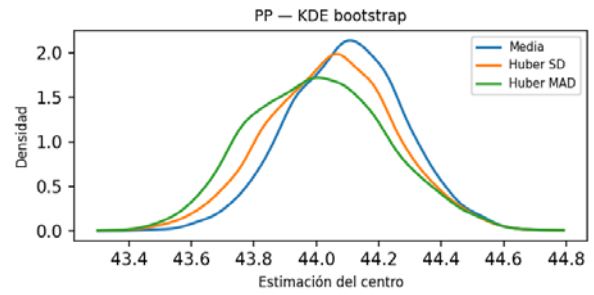
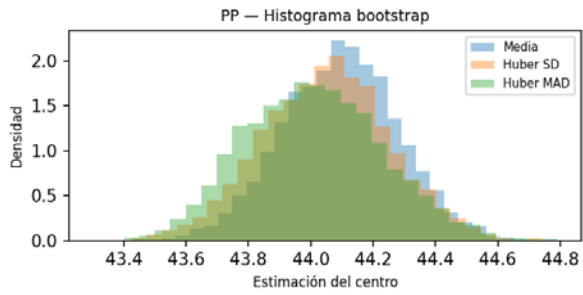
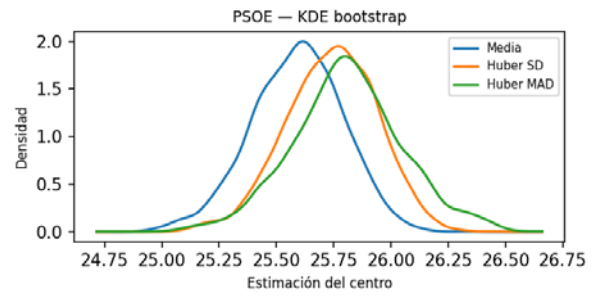
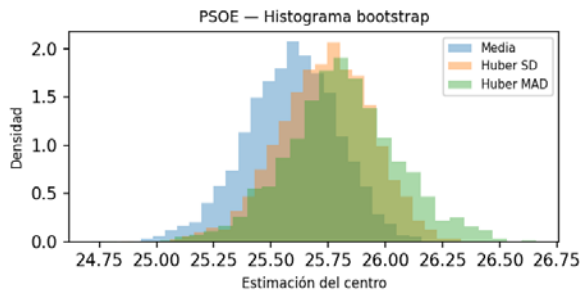
Gráficos con distribuciones bootstrap de estimadores de localización central ( $B = 5000$ , IC 95%). Estos gráficos representan la distribución empírica obtenida mediante bootstrap ( $B = 5000$  réplicas) de distintos estimadores de localización central: media aritmética, Huber con escala basada en desviación estándar (Huber SD) y Huber con escala basada en MAD (Huber MAD). Cada réplica se obtiene re-muestreando escenarios completos, preservando la interdependencia composicional entre partidos y la restricción de suma constante. El histograma muestra la frecuencia relativa de las estimaciones centrales obtenidas en cada réplica, mientras que la curva KDE (Kernel Density Estimation) ofrece una aproximación suavizada de la densidad subyacente. La comparación visual entre estimadores permite evaluar tres dimensiones fundamentales:

- (i) posición relativa del centro,
- (ii) dispersión de la distribución y
- (iii) forma distributiva (asimetría, colas, posible multimodalidad)

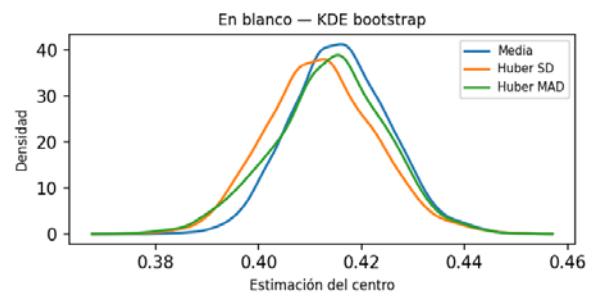
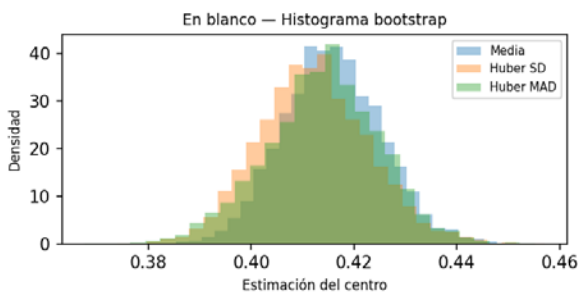
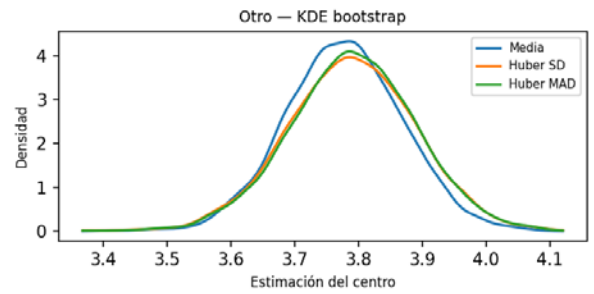
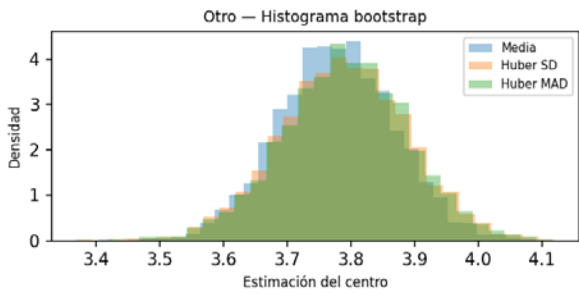
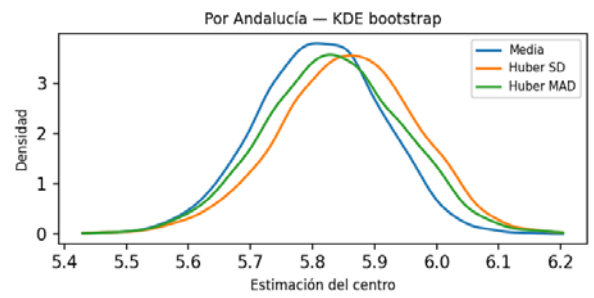
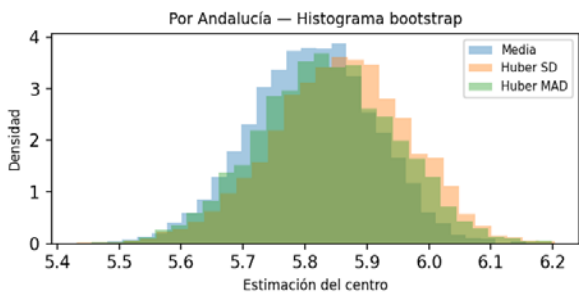
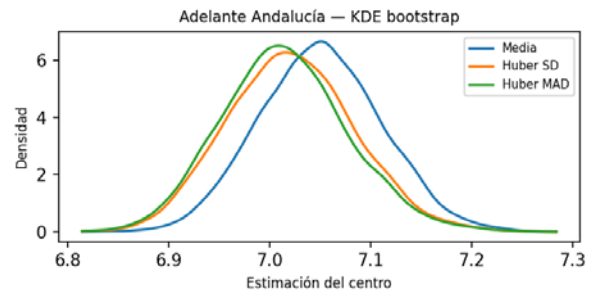
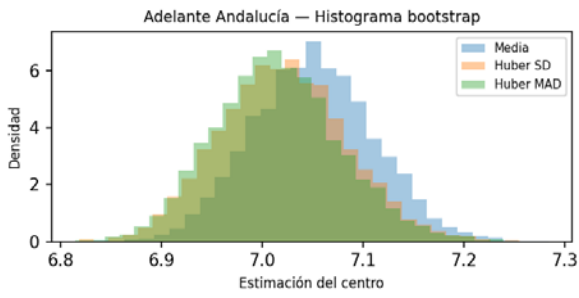
Desde el punto de vista metodológico, la coincidencia sustantiva entre las distribuciones de media y estimadores robustos indica ausencia de valores extremos influyentes y simetría aproximada en la distribución bootstrap. Cuando las curvas se superponen de forma consistente y presentan forma unimodal, puede inferirse estabilidad estructural del estimador bajo perturbaciones del sistema. En cambio, desplazamientos sistemáticos entre media y Huber o diferencias en la anchura de las colas señalarían sensibilidad a asimetrías o a configuraciones extremas de escenarios. La amplitud de la distribución refleja la incertidumbre estructural inducida por la variabilidad entre escenarios completos. Distribuciones estrechas implican alta inercia composicional: pequeñas alteraciones en el conjunto de escenarios no modifican significativamente la posición central. Distribuciones más anchas indican mayor elasticidad del sistema y dependencia del equilibrio global del vector proporcional.

El intervalo de confianza empírico del 95% (percentiles 5%-95%) cuantifica el rango dentro del cual se sitúa la estimación central en la gran mayoría de las réplicas. Si este intervalo es compacto y coherente entre estimadores, la recomendación final se considera robusta. Si el intervalo es amplio o difiere entre estimadores, puede existir tensión estructural o posible transición de régimen probabilístico. En conjunto, estos gráficos permiten distinguir entre fluctuación aritmética y variabilidad composicional real. No solo muestran precisión estadística, sino que verifican la estabilidad del centro estimado cuando se preserva la arquitectura completa del sistema electoral en cada réplica bootstrap.

## PSOE | PP | VOX | SALF



## Adelante Andalucía | Por Andalucía | Otro | En blanco

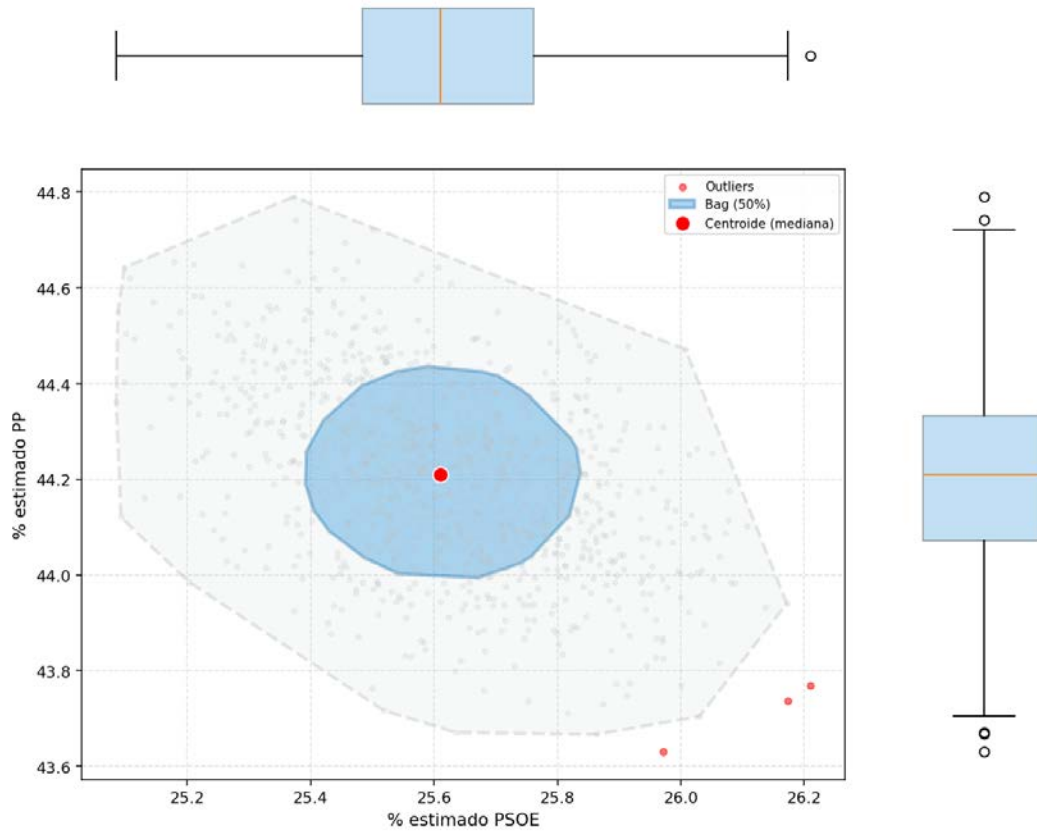


## IV Bagplots bifactoriales entre partidos

En el marco de este modelo, los bagplots trascienden su función como simples herramientas de visualización para constituirse en representaciones rigurosas del espacio de probabilidad inducido por el sistema de escenarios. En lugar de ofrecer una proyección lineal, estos gráficos delimitan la morfología del campo probabilístico, permitiendo observar cómo se distribuye la densidad de las trayectorias analizadas en un plano bidimensional. Esta aproximación es fundamentalmente descriptivo-estructural: no pretende capturar un evento futuro aislado, sino cartografiar la región de configuraciones potenciales que resultan matemáticamente coherentes con la inercia y la incertidumbre inherentes al modelo.

El 'bag' central, al contener el 50% de las observaciones con mayor profundidad estadística, actúa como el núcleo de máxima verosimilitud, mientras que el 'fence' o región extendida define las fronteras de lo plausible, identificando valores atípicos que podrían señalar cambios de régimen o eventos extremos. Al analizar esta geometría, es posible evaluar no solo la tendencia central, sino también la dispersión relativa y las asimetrías que revelan sesgos en el comportamiento del sistema. De este modo, los bagplots se consolidan como dispositivos de diagnóstico que transforman datos multidimensionales en una estructura visual legible, proporcionando un marco de referencia robusto para la toma de decisiones basada en la envolvente de probabilidad del sistema modelizado.

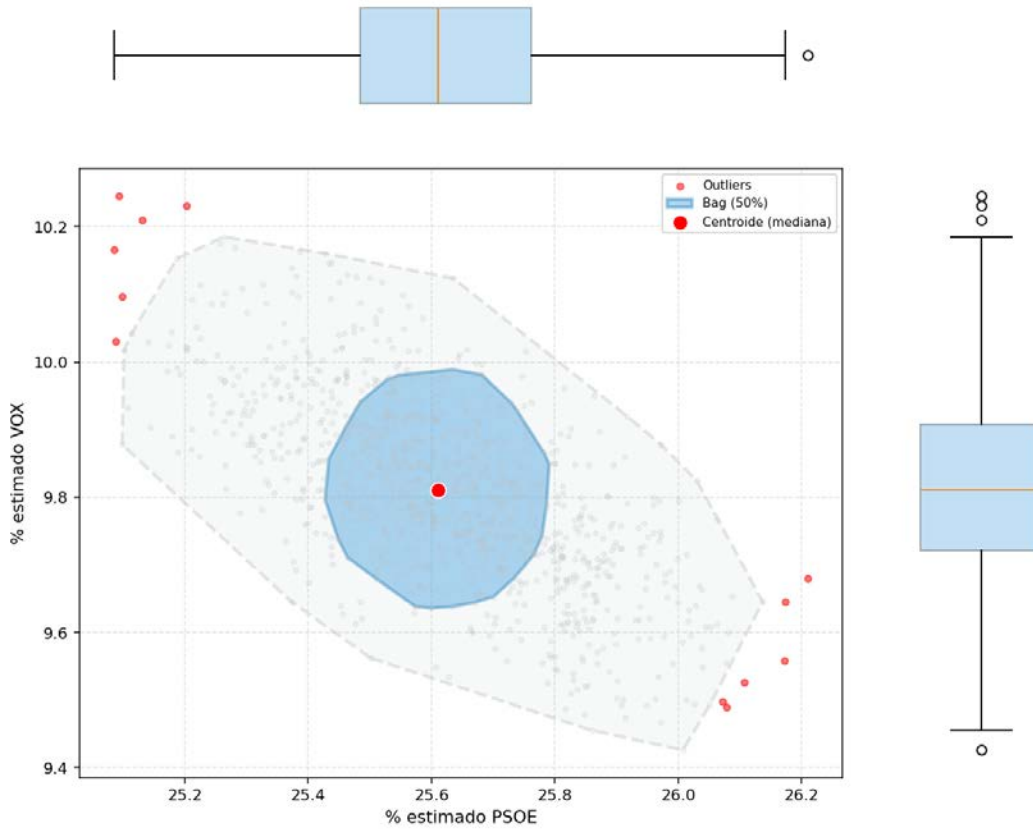
## Covariación bifactorial: PSOE vs PP



### Resumen estadístico bootstrap: PSOE vs PP

Partido	Mediana %	Media %	Desv. Est.	I.C. Inf.	I.C. Sup.
PSOE	25.6	25.6	0.2	25.2	26.0
PP	44.2	44.2	0.2	43.8	44.6

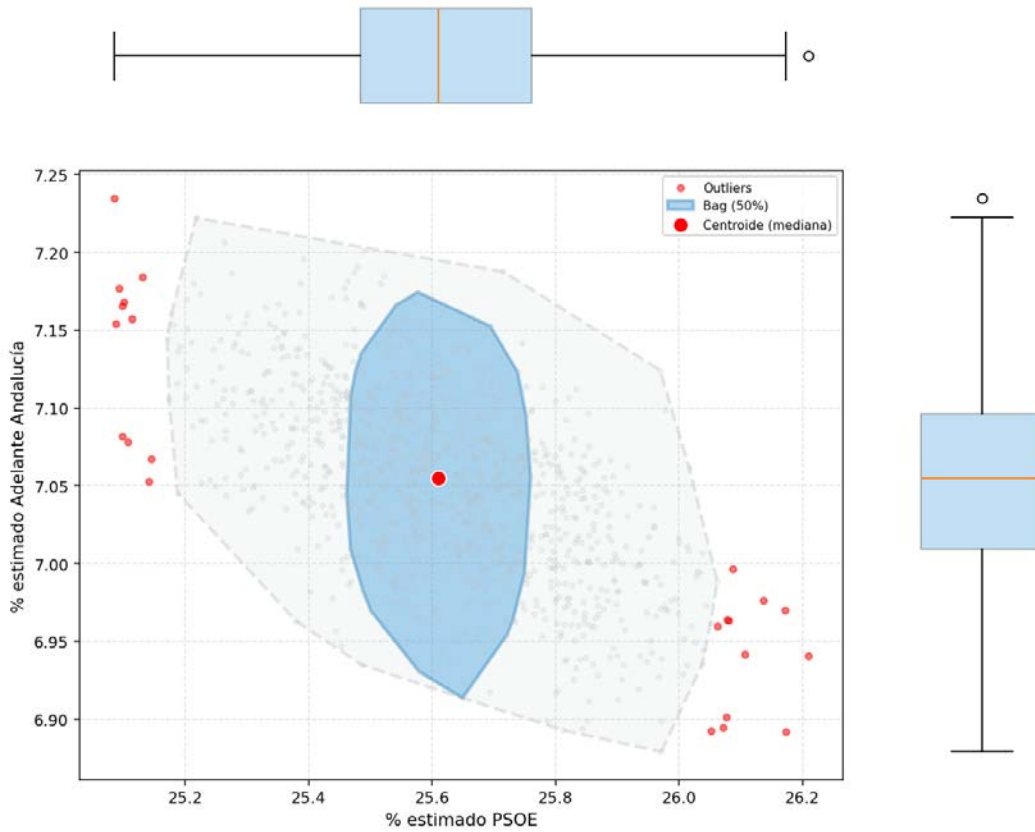
Covariación bifactorial: PSOE vs VOX



Resumen estadístico bootstrap: PSOE vs VOX

Partido	Mediana %	Media %	Desv. Est.	I.C. Inf.	I.C. Sup.
PSOE	25.6	25.6	0.2	25.2	26.0
VOX	9.8	9.8	0.1	9.5	10.1

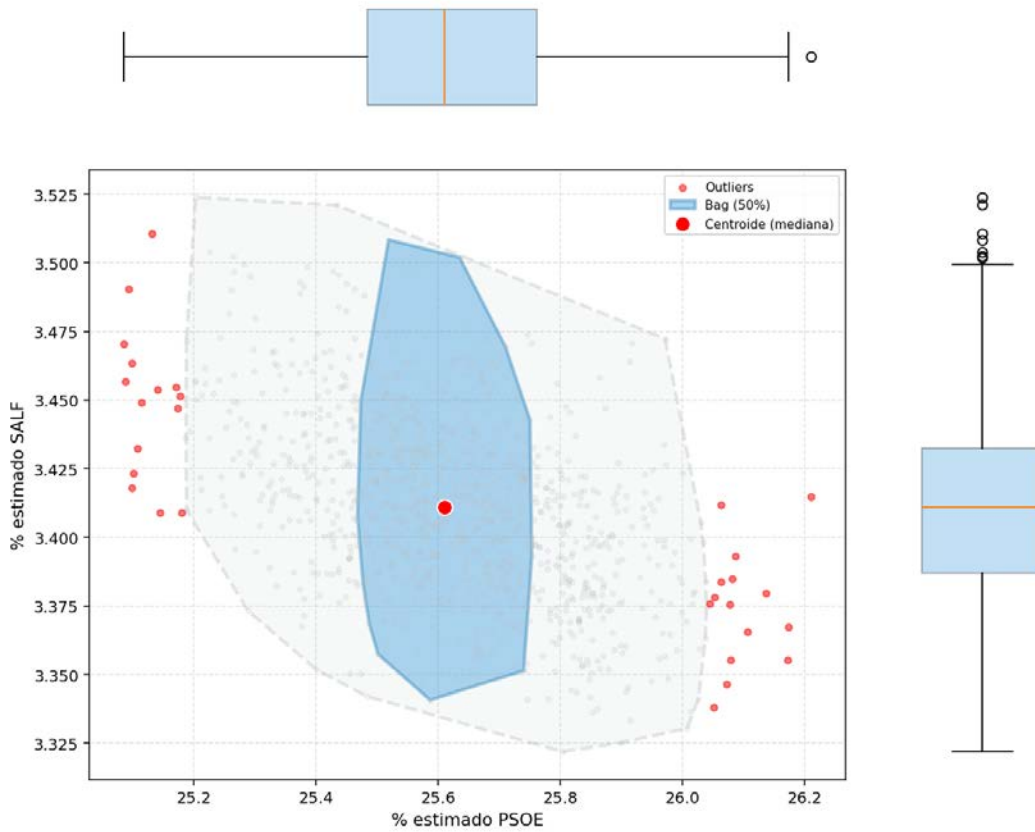
## Covariación bifactorial: PSOE vs Adelante Andalucía



### Resumen estadístico bootstrap: PSOE vs Adelante Andalucía

Partido	Mediana %	Media %	Desv. Est.	I.C. Inf.	I.C. Sup.
PSOE	25.6	25.6	0.2	25.2	26.0
Adelante Andalucía	7.1	7.1	0.1	6.9	7.2

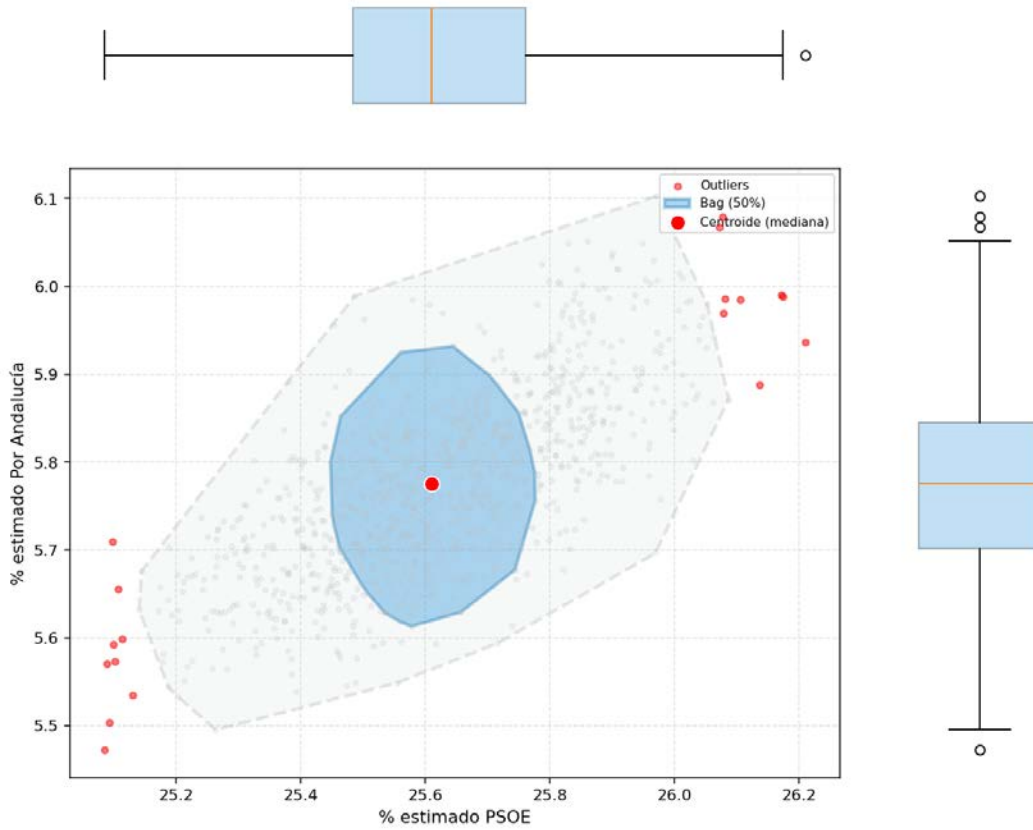
## Covariación bifactorial: PSOE vs SALF



### Resumen estadístico bootstrap: PSOE vs SALF

Partido	Mediana %	Media %	Desv. Est.	I.C. Inf.	I.C. Sup.
PSOE	25.6	25.6	0.2	25.2	26.0
SALF	3.4	3.4	0.0	3.4	3.5

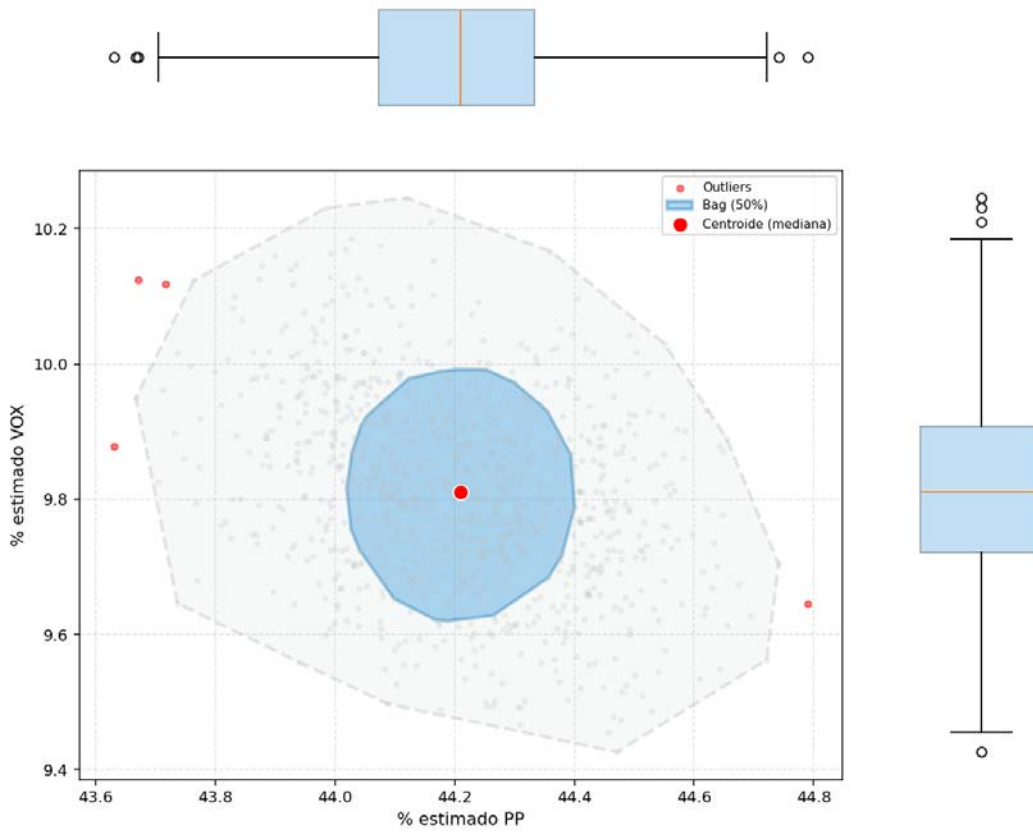
Covariación bifactorial: PSOE vs Por Andalucía



Resumen estadístico bootstrap: PSOE vs Por Andalucía

Partido	Mediana %	Media %	Desv. Est.	I.C. Inf.	I.C. Sup.
PSOE	25.6	25.6	0.2	25.2	26.0
Por Andalucía	5.8	5.8	0.1	5.6	6.0

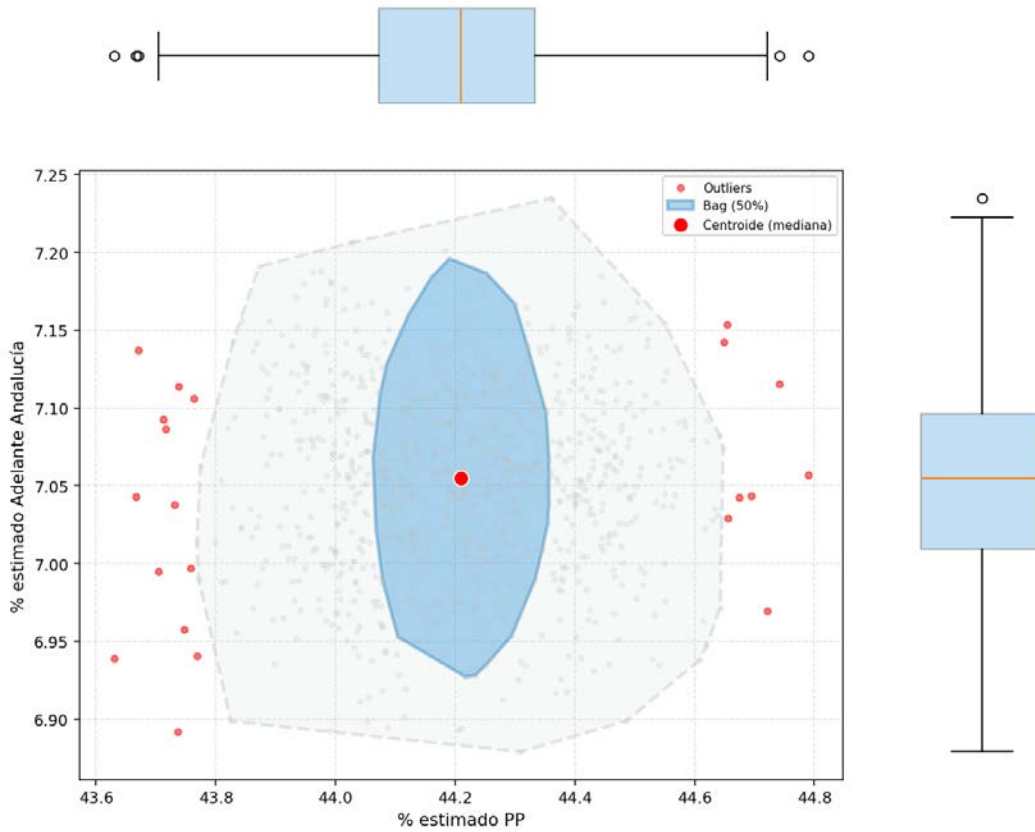
**Covariación bifactorial: PP vs VOX**



**Resumen estadístico bootstrap: PP vs VOX**

Partido	Mediana %	Media %	Desv. Est.	I.C. Inf.	I.C. Sup.
PP	44.2	44.2	0.2	43.8	44.6
VOX	9.8	9.8	0.1	9.5	10.1

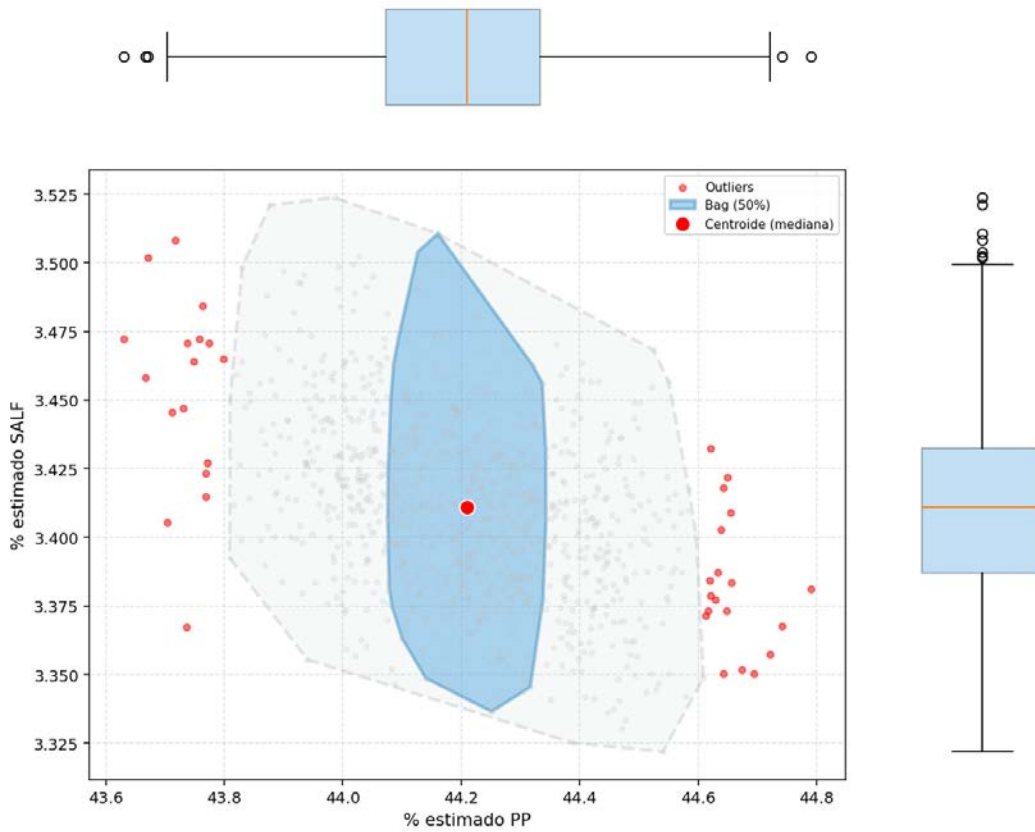
## Covariación bifactorial: PP vs Adelante Andalucía



### Resumen estadístico bootstrap: PP vs Adelante Andalucía

Partido	Mediana %	Media %	Desv. Est.	I.C. Inf.	I.C. Sup.
PP	44.2	44.2	0.2	43.8	44.6
Adelante Andalucía	7.1	7.1	0.1	6.9	7.2

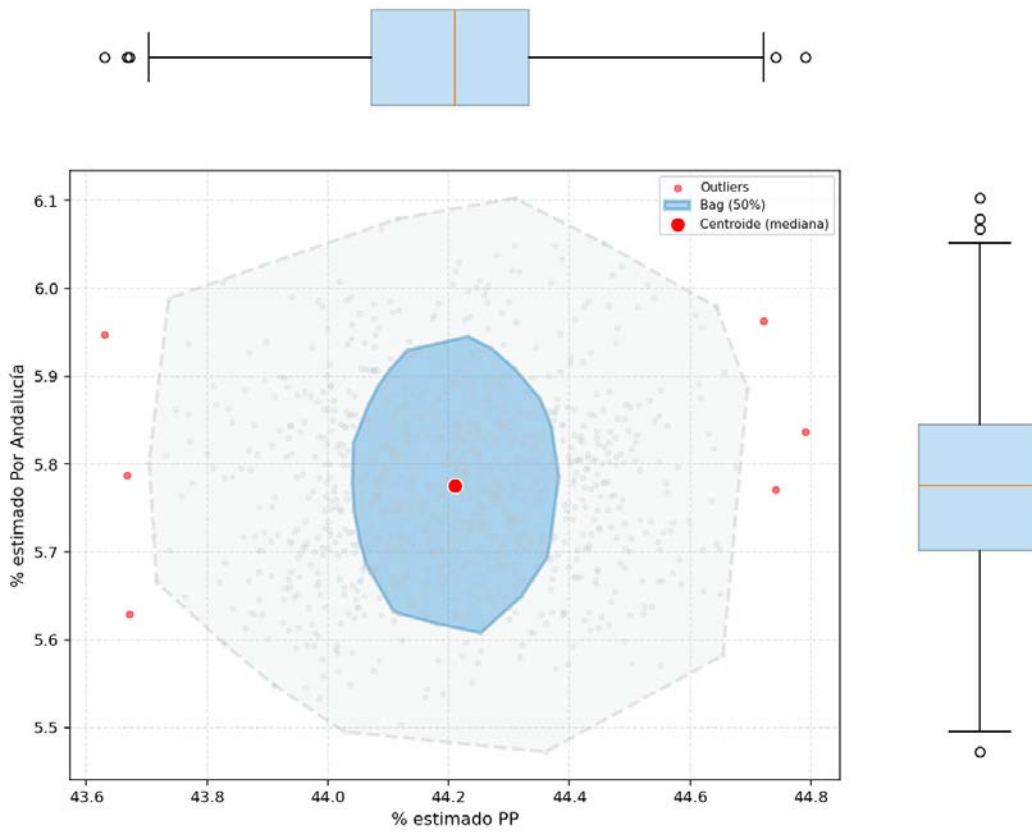
## Covariación bifactorial: PP vs SALF



## Resumen estadístico bootstrap: PP vs SALF

Partido	Mediana %	Media %	Desv. Est.	I.C. Inf.	I.C. Sup.
PP	44.2	44.2	0.2	43.8	44.6
SALF	3.4	3.4	0.0	3.4	3.5

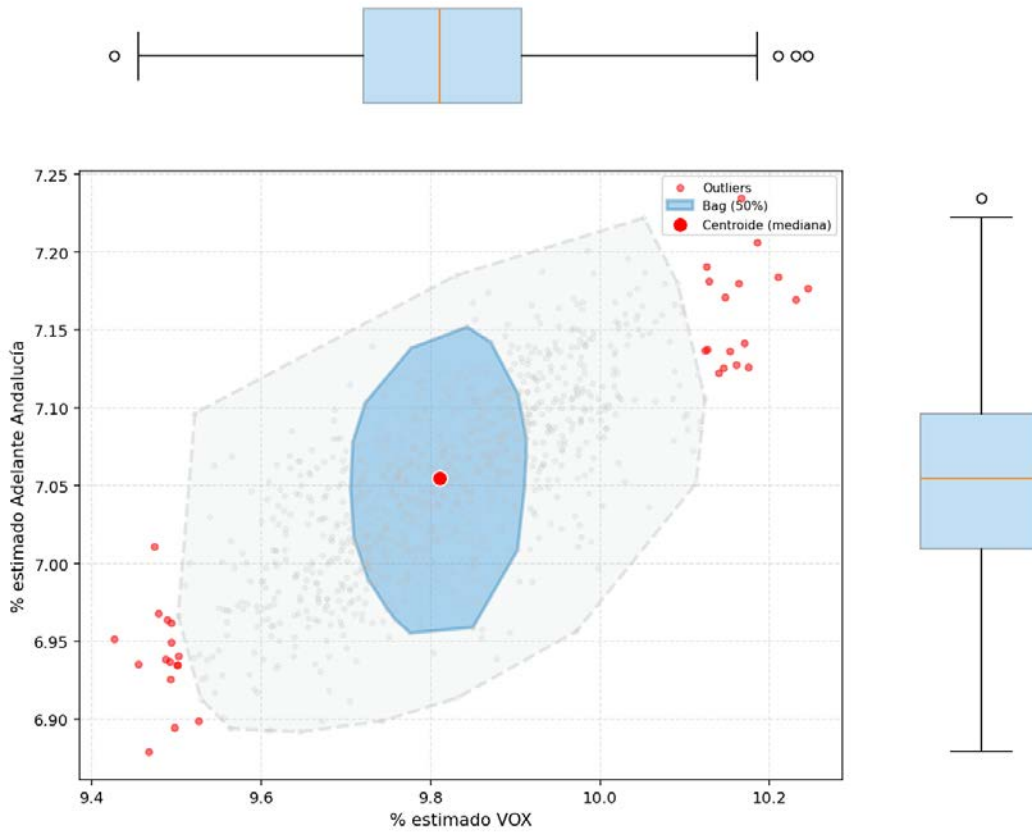
## Covariación bifactorial: PP vs Por Andalucía



### Resumen estadístico bootstrap: PP vs Por Andalucía

Partido	Mediana %	Media %	Desv. Est.	I.C. Inf.	I.C. Sup.
PP	44.2	44.2	0.2	43.8	44.6
Por Andalucía	5.8	5.8	0.1	5.6	6.0

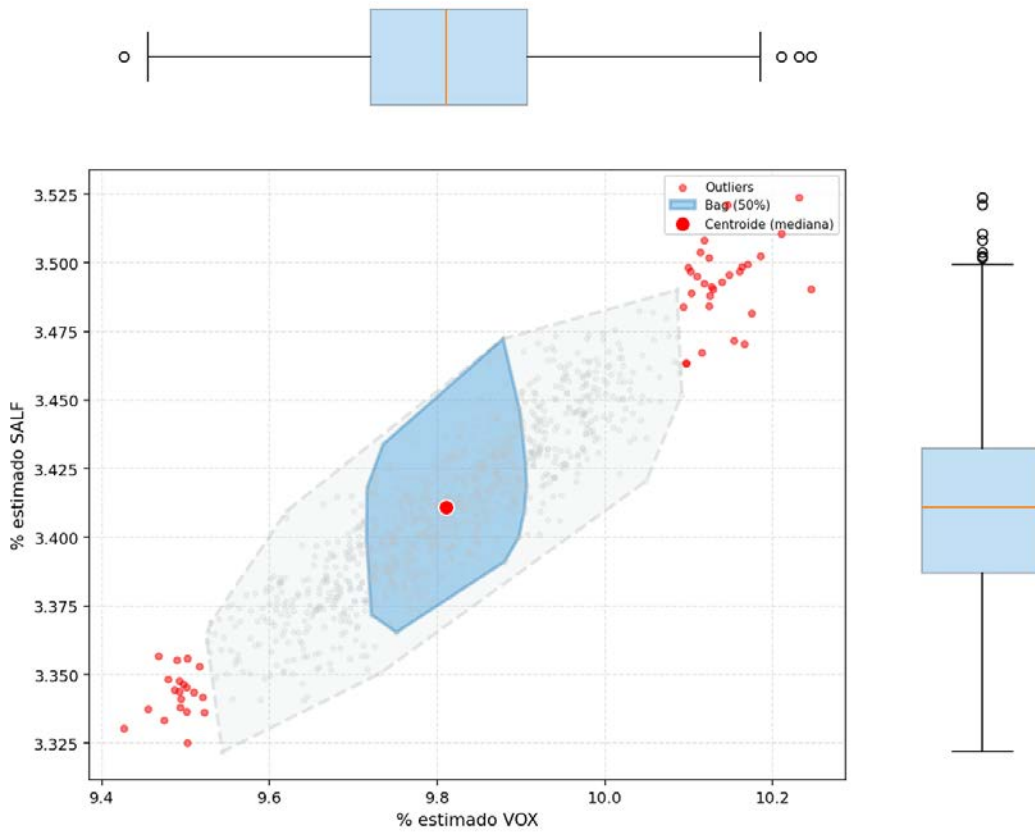
## Covariación bifactorial: VOX vs Adelante Andalucía



### Resumen estadístico bootstrap: VOX vs Adelante Andalucía

Partido	Mediana %	Media %	Desv. Est.	I.C. Inf.	I.C. Sup.
VOX	9.8	9.8	0.1	9.5	10.1
Adelante Andalucía	7.1	7.1	0.1	6.9	7.2

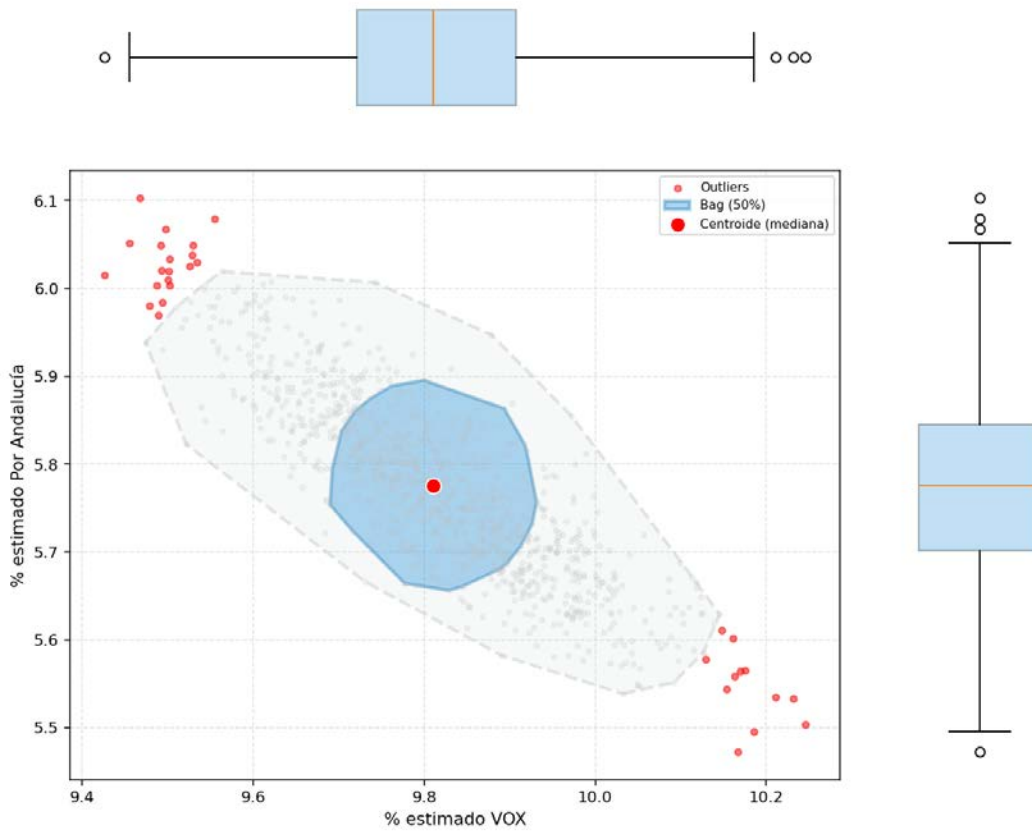
## Covariación bifactorial: VOX vs SALF



### Resumen estadístico bootstrap: VOX vs SALF

Partido	Mediana %	Media %	Desv. Est.	I.C. Inf.	I.C. Sup.
VOX	9.8	9.8	0.1	9.5	10.1
SALF	3.4	3.4	0.0	3.4	3.5

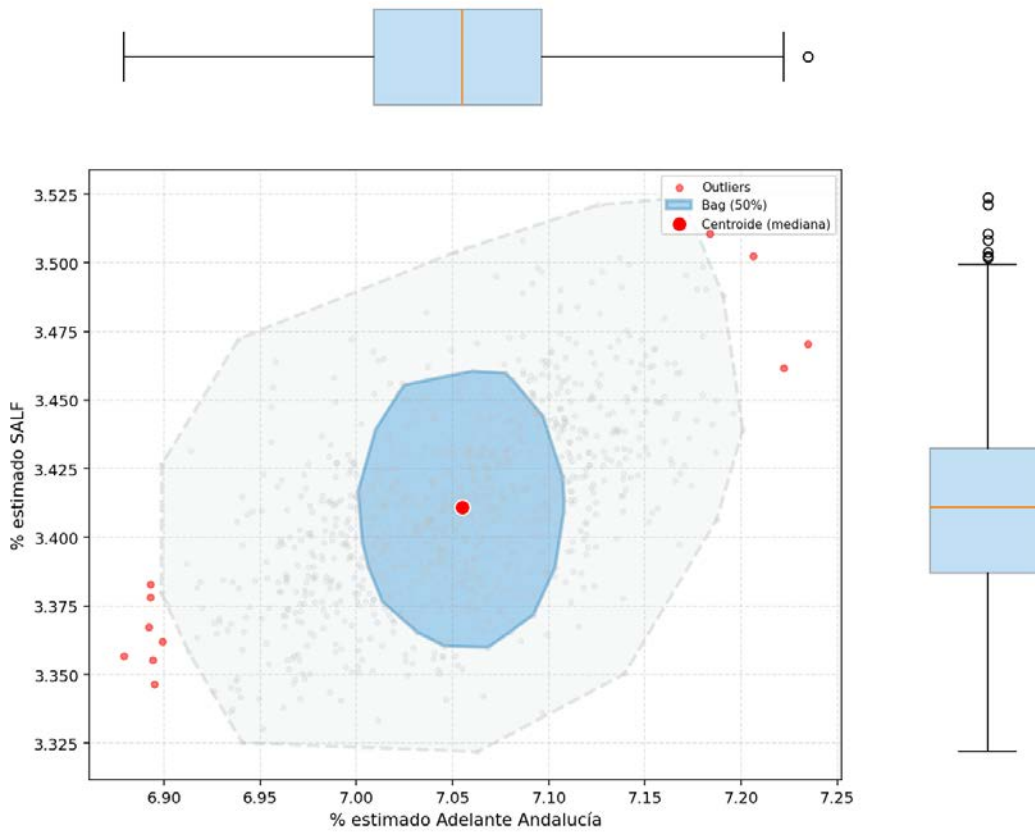
## Covariación bifactorial: VOX vs Por Andalucía



### Resumen estadístico bootstrap: VOX vs Por Andalucía

Partido	Mediana %	Media %	Desv. Est.	I.C. Inf.	I.C. Sup.
VOX	9.8	9.8	0.1	9.5	10.1
Por Andalucía	5.8	5.8	0.1	5.6	6.0

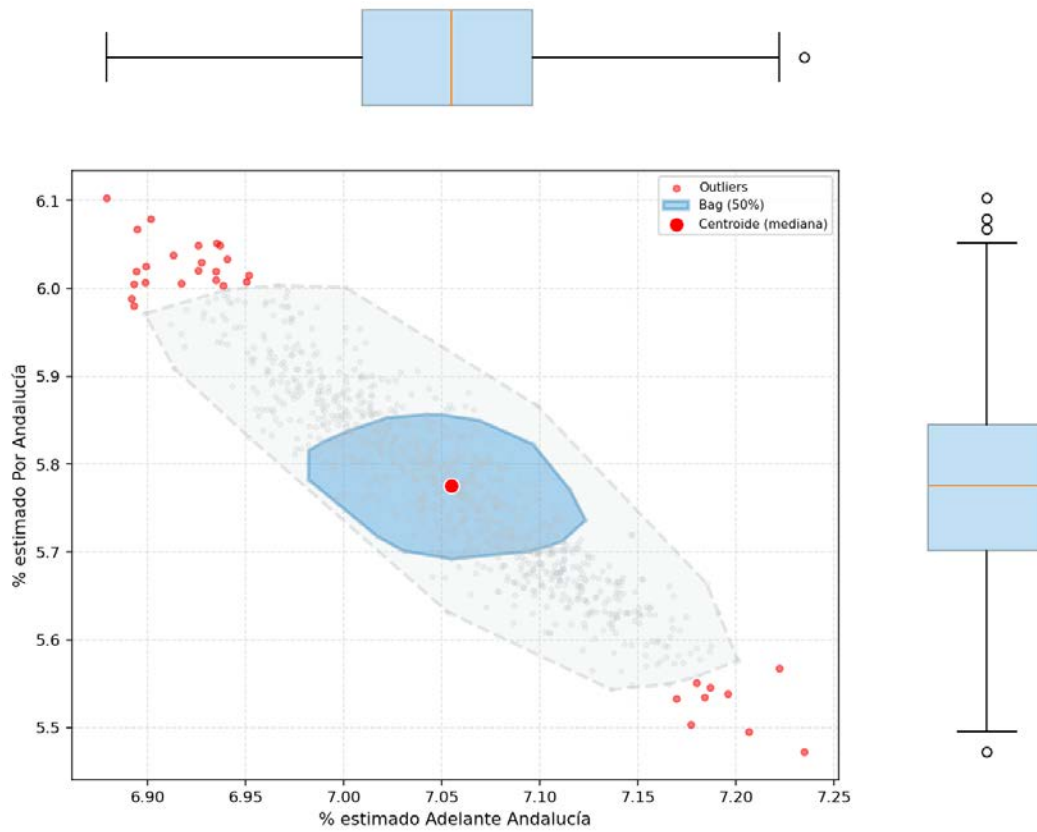
### Covariación bifactorial: Adelante Andalucía vs SALF



### Resumen estadístico bootstrap: Adelante Andalucía vs SALF

Partido	Mediana %	Media %	Desv. Est.	I.C. Inf.	I.C. Sup.
Adelante Andalucía	7.1	7.1	0.1	6.9	7.2
SALF	3.4	3.4	0.0	3.4	3.5

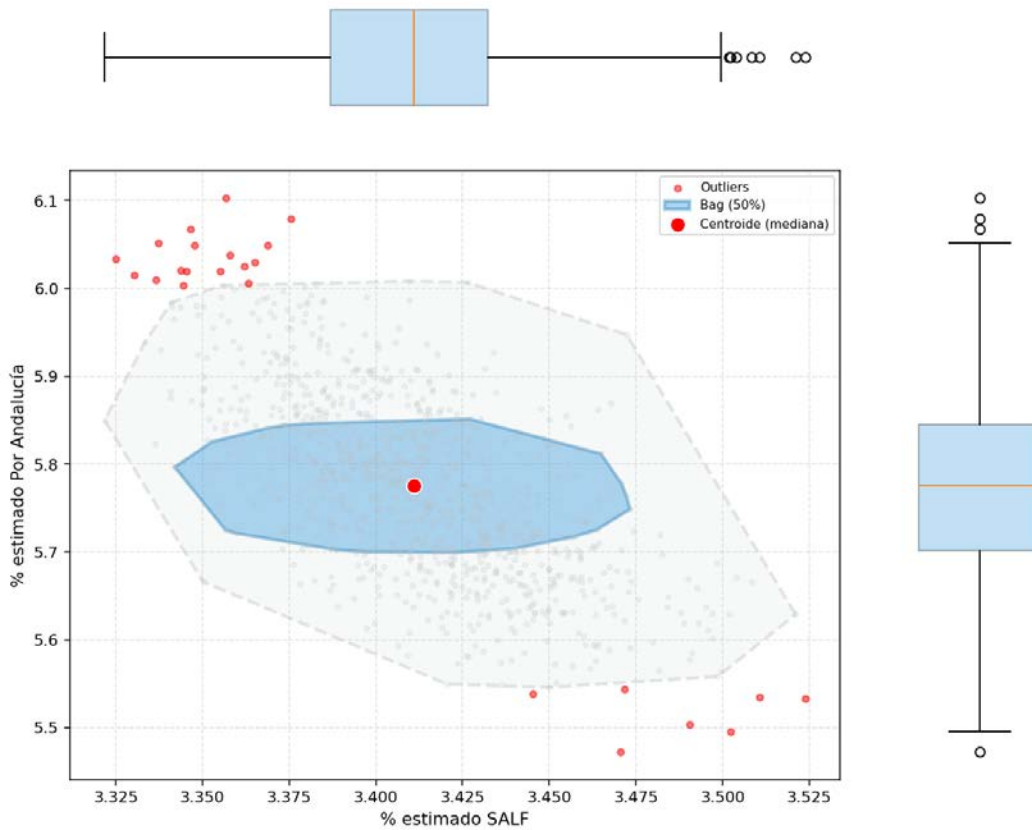
## Covariación bifactorial: Adelante Andalucía vs Por Andalucía



## Resumen estadístico bootstrap: Adelante Andalucía vs Por Andalucía

Partido	Mediana %	Media %	Desv. Est.	I.C. Inf.	I.C. Sup.
Adelante Andalucía	7.1	7.1	0.1	6.9	7.2
Por Andalucía	5.8	5.8	0.1	5.6	6.0

## Covariación bifactorial: SALF vs Por Andalucía



### Resumen estadístico bootstrap: SALF vs Por Andalucía

Partido	Mediana %	Media %	Desv. Est.	I.C. Inf.	I.C. Sup.
SALF	3.4	3.4	0.0	3.4	3.5
Por Andalucía	5.8	5.8	0.1	5.6	6.0

## V. Anexo Técnico: Sistema Integrado de Estimación Robusta y Diagnóstico Estructural

El presente anexo documenta formalmente el sistema implementado para la estimación robusta del centro y el diagnóstico estructural de distribuciones derivadas de escenarios. Todas las formulaciones aquí descritas corresponden exactamente a los cálculos realizados por el programa. El modelo integra estadística robusta, análisis de modalidad, medición de intensidad estructural y geometría composicional.

La arquitectura del sistema se organiza en cuatro planos analíticos: plano aritmético (centro), plano estructural (forma), plano intensivo (segmentación) y plano relacional composicional.

El sistema integra cuatro planos analíticos diferenciados. A continuación, se resume su formulación matemática coherente con la implementación computacional.

### Plano aritmético (criterio decisonal)

$$Eff_{vs\ Mediana}(\theta) = \left( \frac{SE(Mediana)}{SE(\theta)} \right)^2$$

$$Eff^*(\theta) = Eff_{vs\ Mediana}(\theta) \cdot (1 - M)$$

$$\hat{\theta} = \operatorname{argmax}_{\theta \in \Theta} Eff^*(\theta)$$

donde  $\Theta = \{\text{Media, Mediana, Huber\_SD, Huber\_MAD}\}$ . La selección automática del estimador óptimo se realiza exclusivamente en este plano.

### Robustez estructural

$$IRP = \frac{1}{1 + CV_{rob} + S}$$

$$IRP^* = IRP \cdot (1 - M)$$

IRP captura estabilidad intrínseca del centro porcentual; IRP\* incorpora penalización por segmentación estructural.

### Intensidad estructural

$$M = \operatorname{clip} \left[ \frac{BC - 0.555}{(1 - 0.555)}, 0, 1 \right]$$

$M \in [0,1]$ . Valores elevados indican segmentación estructural entre escenarios.

## Plano relacional composicional

$$clr(x_i) = \ln\left(\frac{x_i}{g(x)}\right)$$

$$ilr = \sqrt{\frac{rs}{(r+s)}} \cdot \ln\left(\frac{g(x_R)}{g(x_S)}\right)$$

Las transformaciones CLR e ILR permiten evaluar estabilidad en términos relativos bajo restricción de suma constante. Estas métricas operan como validación estructural y no intervienen en la selección del estimador óptimo.

### A. Plano Aritmético: Estimación Robusta del Centro

Se emplean cuatro estimadores: Media aritmética, Mediana, Huber-SD y Huber-MAD.

#### 1. Media Aritmética

$$\bar{x} = (1/n) \cdot \sum x_i$$

La media presenta eficiencia máxima bajo normalidad, pero función de influencia no acotada.

#### 2. Mediana

$$m = \operatorname{argmin} \sum |x_i - \theta|$$

Posee punto de ruptura del 50% y máxima robustez frente a valores extremos.

### 3. Estimador de Huber (M-estimador)

$$\hat{\mu} = \operatorname{argmin} \sum \rho(x_i - \theta)$$

$$\rho(r) = \frac{1}{2}r^2 \text{ si } |r| \leq c ; c|r| - \frac{1}{2c} \text{ si } |r| > c$$

$$c = 1.345 \cdot s$$

Se implementa mediante iteración ponderada con pesos  $w_i$  dependientes de los residuos truncados.

Con  $c = 1.345 \cdot s$  ( $\approx 95\%$  eficiencia bajo normalidad).

### B. Escalas Robustas

Desviación estándar:

$$s = \sqrt{\left[ \frac{1}{(n-1)} \cdot \sum (x_i - \bar{x})^2 \right]}$$

$$MAD = \operatorname{mediana}(|x_i - \operatorname{mediana}(x)|)$$

### C. Índice de Robustez Penalizada (IRP\*)

El índice IRP cuantifica la estabilidad del centro en el plano porcentual combinando variación robusta relativa y sensibilidad a asimetría. En la implementación, el estimador robusto de referencia es  $\hat{\mu}_{\text{HuberMAD}}$  (Huber con escala MAD).

Escala robusta y variación relativa:

$$\sigma_{rob} = 1.4826 \cdot MAD$$

$$CV_{rob} = \sigma_{rob} / \hat{\mu}_{HuberMAD}$$

Sensibilidad del centro (asimetría relativa):

$$S = | \bar{x} - \hat{\mu}_{HuberMAD} | / \hat{\mu}_{HuberMAD}$$

Índice base de robustez:

$$IRP = 1 / (1 + CV_{rob} + S)$$

La penalización estructural utiliza el score M derivado del coeficiente de bimodalidad (BC), definido como:

$$M = \text{clip}[(BC - 0.555) / (1 - 0.555), 0, 1]$$

BC\_REF = 0.555 corresponde al valor teórico bajo distribución uniforme. La normalización garantiza  $M \in [0,1]$  y permite interpretar M como intensidad estructural relativa.

Robustez penalizada:

$$IRP^* = IRP \cdot (1 - M)$$

$IRP^* \in [0,1]$ . Valores elevados indican simultáneamente: (i) baja variación robusta relativa, (ii) baja asimetría del centro respecto a la media, y (iii) baja segmentación estructural entre escenarios.

## D. Diagnóstico de Forma y Segmentación

$$Dip = \sup |F_n(x) - F_{unimodal}(x)|$$

$$BC = (g_1^2 + 1) / (g_2 + 3)$$

$$M = \text{clip}[(BC - 0.555) / (1 - 0.555), 0, 1]$$

$M \in [0,1]$ . Valores elevados indican segmentación estructural entre escenarios.

## E. Eficiencia Bootstrap y Penalización

$$SE(\theta) = sd\_bootstrap(\theta)$$

$$Eff\_vs\_Mediana(\theta) = (SE(Mediana) / SE(\theta))^2$$

$$Eff^*(\theta) = Eff\_vs\_Mediana(\theta) \cdot (1 - M)$$

La penalización estructural  $(1-M)$  evita seleccionar estimadores aparentemente precisos bajo segmentación intensa.

## F. Regla de Selección del Estimador Óptimo

$$\hat{\theta} = \underset{\theta \in \Theta}{\text{argmax}} Eff^*(\theta)$$

$$\Theta = \{\text{Media, Mediana, Huber\_SD, Huber\_MAD}\}$$

La selección se realiza exclusivamente en el plano aritmético. Las métricas composicionales (CLR/ILR) operan como validación estructural posterior.

## G. Propiedades del Sistema

- Robustez ante valores atípicos.
- Penalización explícita de segmentación estructural.
- Separación entre precisión estadística y estructura distributiva.
- Coherencia con remuestreo bootstrap.
- Regla decisional no ad hoc basada en maximización de  $Eff^*$ .

## H. Esquema Operativo del Sistema de Selección Óptima

El sistema implementa una secuencia analítica estructurada que garantiza coherencia entre forma distributiva, intensidad estructural y estabilidad del centro.

1. Entrada: vectores de escenarios porcentuales (datos cerrados, suma=100).
2. Estimación robusta del centro: Media, Mediana, Huber-SD, Huber-MAD.
3. Diagnóstico estructural:
  - Modalidad formal (Dip).
  - Propensión a segmentación (BC).
  - Intensidad estructural normalizada (M).
4. Penalización estructural de la estabilidad:

$$IRP^* = IRP \cdot (1 - M)$$

5. Evaluación de precisión mediante bootstrap:

$$Eff^* = Eff_{vs\_Mediana}(\theta) \cdot (1 - M)$$

6. Regla automática de selección:

- a) Si  $M > 0.60 \rightarrow$  se prioriza Huber (mezcla intensa).
- b) Si  $M \leq 0.60 \rightarrow$  se selecciona el estimador con mayor  $Eff^*$ .
- c) Si  $M > 0.80 \rightarrow$  se sugiere evaluación mediante modelos de mezcla.

La robustez del sistema deriva de tres mecanismos integrados:

- Separación analítica entre modalidad e intensidad.
- Penalización estructural explícita de estabilidad y eficiencia.
- Regla decisional no ad hoc, basada en umbrales paramétricos definidos ex ante.

En diseños basados en escenarios esta arquitectura evita seleccionar estimadores únicamente por precisión muestral cuando existe fragmentación interna, garantizando coherencia entre forma distributiva y centro recomendado. La selección del incorpora explícitamente diagnóstico estructural e intensidad de segmentación, evitando decisiones sesgadas bajo heterogeneidad distributiva.

La robustez del sistema deriva de la integración secuencial de estos niveles.

Formalmente, el estimador recomendado puede expresarse como:

$$\hat{\theta} = \operatorname{argmax}_{\theta \in \Theta} Eff^*(\theta)$$

$$Eff^*(\theta) = Eff_{vs\ Mediana}(\theta) \cdot (1 - M)$$

$$Eff_{vs\ Mediana}(\theta) = \left( \frac{SE(Mediana)}{SE(\theta)} \right)^2$$

$$\Theta = \{Media, Mediana, Huber_{SD}, Huber_{MAD}\}$$

$M \in [0,1]$  actúa como penalización estructural multiplicativa.

### Síntesis formal del sistema

Plano aritmético (criterio decisional):

$$Eff_{vs\ Mediana}(\theta) = \left( \frac{SE(Mediana)}{SE(\theta)} \right)^2$$

$$Eff^*(\theta) = Eff_{vs\ Mediana}(\theta) \cdot (1 - M)$$

$$\hat{\theta} = \operatorname{argmax}_{\theta \in \Theta} Eff^*(\theta)$$

Robustez estructural:

$$IRP = \frac{1}{1 + CV_{rob} + S}$$

$$IRP^* = IRP \cdot (1 - M)$$

Intensidad estructural:

$$M = clip\left[\frac{BC - 0.555}{(1 - 0.555)}, 0, 1\right]$$

Plano relacional:

$$clr(x_i) = \ln\left(\frac{x_i}{g(x)}\right)$$

$$ilr = \sqrt{\frac{rs}{(r + s)}} \cdot \ln\left(\frac{g(x_R)}{g(x_S)}\right)$$

## Glosario de términos

Se recoge a continuación un glosario operativo de los principales términos técnicos y analíticos empleados en el informe. Su finalidad es facilitar una lectura homogénea de los conceptos metodológicos, geométricos e interpretativos utilizados a lo largo del documento.

**Agregación de bloques.** Procedimiento mediante el cual varios componentes originales de una composición se agrupan en categorías más amplias para facilitar la interpretación.

**Aitchison, distancia de.** Métrica específica para comparar composiciones respetando su naturaleza relativa.

**Análisis composicional.** Enfoque estadístico diseñado para estudiar datos formados por partes de un total.

**Balance.** Contraste log-ratio entre grupos de componentes de una composición.

**Baricentro.** Punto geométrico de equilibrio teórico en una representación composicional.

**Biplot.** Representación gráfica conjunta de observaciones y componentes o cargas.

**Bootstrap.** Procedimiento de re-muestreo utilizado para evaluar la estabilidad de estimadores.

**Centro composicional.** Medida central adaptada a datos composicionales.

**Centro robusto.** Estimación central menos sensible a observaciones extremas.

**Cierre.** Restricción por la cual todas las partes de una composición suman una constante.

**CLR (centered log-ratio).** Transformación log-ratio centrada que expresa cada componente respecto a la media geométrica.

**Componente.** Cada una de las partes que integran una composición.

**Composición.** Vector de partes positivas que expresan un reparto relativo de un total común.

**Contribución a la distancia.** Participación relativa de cada componente en la distancia composicional total entre dos escenarios.

**Descomposición de la distancia.** Procedimiento que reparte la distancia total entre componentes.

**Desplazamiento composicional.** Cambio de posición de una observación en el espacio composicional.

**Diagrama ternario.** Representación gráfica de composiciones de tres partes en un simplex triangular.

**Dispersión.** Grado de separación entre observaciones dentro del espacio analítico.

**Escenario.** Configuración analítica que representa un estado, referencia o proyección del sistema.

- Espacio composicional.** Espacio matemático en el que se interpretan las composiciones y sus relaciones.
- Geometría de Aitchison.** Marco geométrico específico para el análisis de composiciones.
- Heatmap.** Representación matricial en la que los valores se expresan mediante una escala de color.
- Hito.** Observación destacada dentro de una trayectoria o comparación.
- ILR (isometric log-ratio).** Transformación log-ratio isométrica que produce coordenadas ortonormales.
- Media geométrica.** Medida central multiplicativa utilizada en análisis composicional.
- Mediana.** Valor central de una distribución, útil como referencia robusta.
- Observación.** Unidad individual analizada dentro del informe.
- PCA (principal component analysis).** Técnica de reducción dimensional que resume la variación principal del sistema.
- Plano ILR.** Representación bivariada construida a partir de coordenadas ILR.
- Re-cierre.** Normalización de una composición para que sus componentes vuelvan a sumar una constante.
- Reducción dimensional.** Proceso mediante el cual un conjunto amplio de variables se resume en menos dimensiones.
- Robustez.** Capacidad de un resultado para mantenerse estable ante pequeñas perturbaciones.
- Simplex.** Espacio geométrico natural de las composiciones.
- Transformación.** Operación matemática aplicada a los datos para analizarlos en un marco adecuado.
- Trayectoria.** Secuencia de posiciones adoptadas por observaciones ordenadas.
- Validación.** Conjunto de procedimientos destinados a valorar la solidez de los resultados.
- Variación.** Cambio observado entre componentes, escenarios o posiciones.
- Vector de carga.** Dirección y peso de un componente en un biplot o PCA.

## V Referencias Metodológicas

- Aitchison, J. (1982). The statistical analysis of compositional data (with discussion). *Journal of the Royal Statistical Society, Series B (Statistical Methodology)*, 44(2), 139–177.
- Aitchison, J. (1986). *The Statistical Analysis of Compositional Data*. Monographs on Statistics and Applied Probability. Chapman & Hall, London. (Reprinted 2003, Blackburn Press).
- Aitchison, J. (1997). The one-hour course in compositional data analysis or compositional data analysis is simple. In V. Pawlowsky-Glahn (Ed.), *Proceedings of iAMG'97 – The Third Annual Conference of the International Association for Mathematical Geology* (pp. 3–35). CIMNE, Barcelona.
- Aitchison, J., Barceló-Vidal, C., Egozcue, J. J., & Pawlowsky-Glahn, V. (2002). A concise guide for the algebraic-geometric structure of the simplex, the sample space for compositional data analysis. In U. Bayer, H. Burger, & W. Skala (Eds.), *Proceedings of iAMG'02* (pp. 387–392). Alfred-Wegener-Stiftung, Berlin.
- Aitchison, J., & Greenacre, M. (2002). Biplots for compositional data. *Journal of the Royal Statistical Society, Series C (Applied Statistics)*, 51(4), 375–392.
- Aitchison, J., & Shen, S. M. (1980). Logistic-normal distributions: Some properties and uses. *Biometrika*, 67(2), 261–272.
- Alaminos, A., & Tezanos, J. F. (2019a). Nota metodológica estimaciones de voto y escaños. Modelo coyuntura CIS V89. Estudio nº 3261. Barómetro de septiembre 2019.
- Alaminos, A., & Tezanos, J. F. (2019b). Nota metodológica estimaciones de voto y escaños. Estudio nº 3263 Macrobarómetro Preelectoral elecciones generales 2019.
- Alaminos, A. (2024). Introducción a la medición de apoyos electorales mediante el método de escenarios. Modelo bifactorial Inercia-Incertidumbre. Alicante: Obets Ciencia Abierta.
- Alaminos, A., & Alaminos-Fernández, A. F. (2026a). *Geometría de la Competencia Política. Una teoría composicional del sistema de partidos*. Alicante: Obets Ciencia Abierta.
- Alaminos, A., & Alaminos-Fernández, A. F. (2026b). *Introducción al análisis de datos composicionales (CoDA)*. Alicante: Obets Ciencia Abierta.
- Alaminos, A., & Alaminos-Fernández, A. F. (2024). Análisis de las dinámicas electorales en las elecciones autonómicas de Galicia del 18 de febrero de 2024. Modelo Bifactorial Inercia-Incertidumbre. Número 81. Opiniones y Actitudes. Madrid: CIS.
- Alaminos, A., & Alaminos-Fernández, A. F. (2023a). *Modelo Bifactorial Inercia-Incertidumbre Alaminos-Tezanos. Aplicación al diagnóstico preelectoral y evaluación del impacto de campaña*. Colección Cuadernos Metodológicos (Vol. 63). Madrid: Centro de Investigaciones Sociológicas.
- Alaminos, A., & Alaminos-Fernández, A. F. (2023b). Motivaciones de voto y estimación electoral en las elecciones generales en España. Un análisis empírico. *Revista Sistema*, 267, 3–37.

- Alaminos, A. (2022). El método de escenarios en la estimación de resultados electorales. Una aplicación al caso de Castilla y León. *Revista Española de Investigaciones Sociológicas*, 178, 173–182.
- Alaminos, A. (2021). Las dinámicas de cambio en el comportamiento del electorado en España. *Sistema: Revista de Ciencias Sociales*, 260, 17–34.
- Billheimer, D., Guttorp, P., & Fagan, W. F. (2001). Statistical interpretation of species composition. *Journal of the American Statistical Association*, 96(456), 1205–1214.
- Chayes, F. (1960). On correlation between variables of constant sum. *Journal of Geophysical Research*, 65(12), 4185–4193.
- Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, 7(1), 1–26.
- Egozcue, J. J., & Pawlowsky-Glahn, V. (2005). Groups of parts and their balances in compositional data analysis. *Mathematical Geology*, 37(7), 795–828.
- Egozcue, J. J., Pawlowsky-Glahn, V., Mateu-Figueras, G., & Barceló-Vidal, C. (2003). Isometric logratio transformations for compositional data analysis. *Mathematical Geology*, 35(3), 279–300.
- Hartigan, J. A., & Hartigan, P. M. (1985). The Dip Test of Unimodality. *The Annals of Statistics*, 13(1), 70–84.
- Huber, P. J. (1964). Robust estimation of a location parameter. *The Annals of Mathematical Statistics*, 35(1), 73–101.
- Huber, P. J. (1981). *Robust Statistics*. Wiley, New York.
- Martín-Fernández, J. A., Barceló-Vidal, C., & Pawlowsky-Glahn, V. (2003). Dealing with zeros and missing values in compositional data sets using nonparametric imputation. *Mathematical Geology*, 35(3), 253–278.
- Pawlowsky-Glahn, V., & Egozcue, J. J. (2002). BLU estimators and compositional data. *Mathematical Geology*, 34(3), 259–274.
- Rousseeuw, P. J., & Croux, C. (1993). Alternatives to the Median Absolute Deviation. *Journal of the American Statistical Association*, 88(424), 1273–1283.
- Sarle, W. S. (1987). Bimodality coefficient. SAS Technical Report.
- Smithson, M., & Broomell, S. B. (2022). Compositional data analysis tutorial for social science researchers: Techniques for dealing with percentage and constrained-sum data. *Psychological Methods*.  
<https://doi.org/10.1037/met0000464>