



Papeles de Energía

Inversiones sostenibles en tiempos revueltos

Nº31

Febrero 2026

¿Han olvidado los mercados de capitales la sostenibilidad?

Robert Engle

Understanding the Dispersion of Environmental Scores across Industrial Sectors

Juan Ignacio Peña, Kurt A. Desender and Bing Guo

Invertir en la transición energética bajo incertidumbre profunda: modelos de planificación y métodos robustos

Antonio F. Rodríguez Matas

Índice

- 3 **Introducción: Inversiones sostenibles en tiempos revueltos**
- 7 **¿Han olvidado los mercados de capitales la sostenibilidad?**
Robert Engle
- 23 **Understanding the Dispersion of Environmental Scores across
Industrial Sectors**
Juan Ignacio Peña, Kurt A. Desender and Bing Guo
- 43 **Invertir en la transición energética bajo incertidumbre profunda:
modelos de planificación y métodos robustos**
Antonio F. Rodríguez Matas

INTRODUCCIÓN

Inversiones sostenibles en tiempos revueltos

Vivimos tiempos convulsos para la transición energética, la descarbonización o, en general, la incorporación de la sostenibilidad en el ámbito empresarial o gubernamental. Las decisiones recientes del presidente Trump que tratan de desmontar las políticas climáticas de sus predecesores, o la creciente incertidumbre acerca del mantenimiento de la ambición climática en Europa, que se está traduciendo en manifestaciones de algunos líderes nacionales a favor de relajar los objetivos e instrumentos climáticos en la Unión, o el cada vez mayor cuestionamiento de los estándares ESG, o de la inversión sostenible, en el ámbito empresarial, son síntomas de un cambio de orientación en la agenda ambiental y climática que los países desarrollados llevaban impulsando desde hacía décadas. Un cambio que viene apoyado en el creciente descontento de una parte de la opinión pública, y canalizado y a veces alimentado por muchos partidos populistas.

En este contexto, cabe preguntarse por el futuro de los marcos de inversión sostenible que se han venido utilizando hasta ahora, y que ya se han cubierto en ocasiones en esta revista. Pero también puede aprovecharse este momento para tratar de establecer sobre bases más sólidas los distintos esquemas de valoración empresarial o los métodos de planificación de inversiones en este nuevo marco, de forma que el retorno hacia la senda deseada sea más robusto. Ese es el objetivo de este número.

En el primer artículo, **Robert Engle**, profesor de la Universidad de Nueva York, y galardonado en 2003 con el premio del Banco de Suecia en Ciencias Económicas en memoria de Alfred Nobel, se pregunta si los mercados de capitales han olvidado la sostenibilidad, a la vista del comportamiento bursátil de la industria fósil y del desempeño de los fondos sostenibles. Para ello introduce el concepto de riesgo de terminación: el riesgo de que una actividad deba cesar, incluso contando con flujos de caja positivos.

Este riesgo de terminación le permite explicar por qué el sector fósil ha tenido un buen desempeño bursátil: el riesgo incentiva a las empresas a reducir su inversión, y por ello a aumentar el flujo de caja. Si el flujo de caja aumenta significativamente, el precio de las acciones también lo hará. Si además se produce un *shock* de demanda (como el posterior a la pandemia), los precios de los fósiles aumentan y las acciones siguen subiendo. Engle recuerda además que este efecto se amplifica al no contar con un impuesto a las emisiones de CO₂: los retornos van a los accionistas en lugar de al fisco.

Ahora bien, el autor considera que sigue teniendo sentido invertir en carteras que cubran el riesgo de terminación. Si se sigue tratando de luchar contra el cambio climático, los cambios en el riesgo de terminación llevarán a las carteras en la dirección correcta, y las acciones de fósiles caerán y tenderán a cero. Pero eso requiere que, también para reducir el riesgo de terminación de la humanidad, los principales países emisores colaboren, cooperen y trabajen juntos hacia ese objetivo.

En el segundo trabajo, **Juan Ignacio Peña, Kurt A. Desender y Bing Guo**, de la Universidad Carlos III de Madrid, analizan la robustez de los indicadores de desempeño ambiental (la E, dentro del marco ESG) y en qué medida su dispersión puede aportar información. Los autores comienzan observando que, en muchas ocasiones, hay desacuerdo entre las empresas de *rating* (dispersión externa), o incluso dentro de una misma valoración (dispersión interna), acerca del desempeño ambiental de las empresas. Y se preguntan por el significado de este desacuerdo.

A partir de datos de EE. UU. en el período 2002-2019, encuentran que la dispersión externa es mayor en las industrias más emisoras, mientras que en las industrias verdes las diferencias se deben más a cuestiones de transparencia y vigilancia institucional. En cambio, la dispersión interna es mayor en las empresas verdes. Ello los lleva a las siguientes conclusiones: en primer lugar, la dispersión no es aleatoria, sino que refleja cuestiones estructurales de las empresas y las agencias de *rating*, lo que sugiere diferencias metodológicas más que la existencia de visiones radicalmente diferentes acerca del comportamiento ambiental de la empresa. En segundo lugar, la dispersión interna muestra fundamentalmente la complejidad de cumplir con todos los requisitos ambientales por parte de las empresas. Tercero, la estructura de propiedad explica también los resultados, a través de las reglas de “disclosure”.

Peña *et al.* recomiendan a los reguladores centrarse en mejorar la transparencia y la coherencia entre las agencias de *rating*, para lograr que los inversores no se vean confundidos por diferencias metodológicas evitables. También tienen un mensaje para los inversores: una única valoración ESG no puede contar la historia completa, es fundamental entender las diferencias entre *ratings* para asignar correctamente las inversiones.

Finalmente, **Antonio F. Rodríguez Matas**, investigador del IIT en la Universidad Pontificia Comillas, presenta distintas herramientas para planificar inversiones de forma robusta frente a tanta incertidumbre actual, en particular en el sector energético, en el que las inversiones son intensivas en capital y de larga vida útil, y en el que es necesario evaluar las inversiones de forma sistémica y no aislada.

El autor defiende que en planificación de inversiones energéticas es preciso tener en cuenta la incertidumbre “profunda” o “epistémica”, es decir, aquella que no se puede representar mediante funciones de probabilidad. Y que para ello no basta trabajar con escenarios “pro-

bables” y contextos de optimización. Las preguntas dejan de ser cuáles son las inversiones óptimas, y pasan a ser qué estrategias siguen funcionando cuando el futuro cambia, o qué incertidumbres hay que vigilar para evitar el colapso de una estrategia.

Para hacer operativas estas ideas, Rodríguez Matas introduce tres herramientas, incorporadas en el modelo de acceso abierto openMASTER: la optimización robusta unida a la minimización del máximo arrepentimiento; la integración de las correlaciones entre parámetros inciertos; y el diseño de políticas robustas. Y propone una guía práctica para incorporar estas herramientas en las decisiones de inversión, que puede ser útil para distintos actores: administraciones públicas, reguladores, empresas o la sociedad civil. También ofrece algunas implicaciones generales para la agenda inversora en la transición energética: priorizar inversiones habilitadoras y de flexibilidad (como las redes, el almacenamiento o la gestión de la demanda), incorporar la demanda como recurso o utilizar los resultados del modelado para monitorizar y adaptar las estrategias.

Confío, como en otras ocasiones, en que estas lecturas sean de utilidad para mantener la senda de la transición energética y acomodar las inversiones necesarias a estos tiempos inciertos.

¿Han olvidado los mercados de capitales la sostenibilidad?

Robert Engle*

Resumen

Este artículo analiza cómo el riesgo climático puede integrarse en la teoría moderna de carteras, introduciendo el concepto de *riesgo de terminación*. Se argumenta que los mercados financieros, al ser prospectivos, incorporan los riesgos climáticos de largo plazo en los precios actuales de los activos.

Palabras clave: riesgo climático, riesgo de terminación, mercados financieros.

1. INTRODUCCIÓN

Celebramos la extraordinaria carrera de Harry Markowitz. De forma sorprendente, casi toda la investigación contemporánea en carteras comienza con la intuición básica de Markowitz. Él nos mostró que las finanzas tratan tanto del riesgo como del retorno, y esta idea ha moldeado la manera en que todos pensamos hoy. En su momento fue revolucionaria y ahora la conocemos como teoría moderna de carteras. Fue publicada en 1952, por lo que en realidad no es tan moderna, pero, en el contexto del arte moderno o la música clásica, se encuentra dentro del rango adecuado.

Me centraré en el riesgo climático, un tema que Markowitz no abordó en su artículo original, pero que constituye una aplicación directa de muchas de sus ideas. Analizaré la noción de riesgo de largo plazo en modelos climáticos e introduciré un tipo particular de riesgo de largo plazo denominado *riesgo de terminación*, que resulta apropiado para comprender el comportamiento económico de empresas expuestas a riesgos climáticos severos.

Markowitz reconoció que cuando utilizamos parámetros estimados de medias y de la matriz de covarianzas, las carteras no presentan el desempeño deseado. Existen múltiples soluciones a este problema, como emplear muestras extremadamente largas o restringir los modelos para reducir el número de parámetros. Un dilema específico del marco de Markowitz que me resulta especialmente interesante es el siguiente: cuando descubrimos que las volatilidades son predecibles, surge naturalmente la pregunta de cómo utilizar esta información dentro del marco de Markowitz para optimizar riesgo y retorno.

* New York University.

¿Han olvidado los mercados de capitales la sostenibilidad?

Puedo afirmar, tanto por mi propio trabajo como por los comentarios de otros investigadores, que si se utilizan datos de alta frecuencia para medir el riesgo y datos de baja frecuencia para medir los retornos, y se combinan directamente, el enfoque no funciona. Esto se debe a que los retornos esperados se mueven al mismo tiempo que las volatilidades, por lo que se están utilizando retornos esperados incorrectos, lo que conduce a un desempeño insatisfactorio. En consecuencia, terminamos utilizando nociones de retornos esperados basadas en datos de largo plazo y nociones de riesgo basadas en datos de alta frecuencia, existiendo un componente de arte en la gestión de esta combinación.

Me interesa particularmente la cuestión de qué ocurre cuando el riesgo cambia. Markowitz no responde a esta pregunta de manera dinámica, en el sentido de analizar cómo se ajustan las carteras cuando el riesgo pasa de un nivel a otro. Cuando el riesgo aumenta, el activo suele volverse menos atractivo y su precio tiende a caer. A medida que el precio cae, el retorno esperado aumenta —suponiendo que los flujos de caja no se vean demasiado afectados— hasta que el precio disminuye lo suficiente como para inducir a los agentes económicos a mantener el activo. De este modo, surge una prima de riesgo más elevada como respuesta al mayor riesgo.

Este proceso ocurre constantemente. Puede tomar minutos o incluso menos, pero es muy frecuente y lo observamos cada vez que hay un cambio en el riesgo. Incluso si el riesgo se materializa en un futuro lejano, creemos que se refleja en los datos actuales, ya que los mercados son prospectivos (*forward looking*). Si esto no resulta convincente, basta observar lo ocurrido con Nvidia y Tesla. Ambas acciones tuvieron valoraciones extremadamente altas basadas en expectativas optimistas sobre el futuro. Sin embargo, el optimismo sobre Tesla se ha reducido y la acción cayó un 27 % de enero a junio de 2024, mientras que la cotización de Nvidia subió casi un 100 % durante el mismo periodo. Esto ilustra cómo los cambios en el riesgo se reflejan en los precios actuales¹.

Si deseamos implementar el marco de Markowitz en el contexto del cambio climático, debemos preguntarnos qué medidas de riesgo utilizar. ¿Debemos emplear medidas históricas o medidas prospectivas? La respuesta es que debemos utilizar medidas prospectivas en la medida de lo posible. Robert Merton proporcionó un marco conceptual para ello hace aproximadamente 50 años. Los inversores pueden anticipar que el futuro será diferente del presente y tratar de cubrir esos cambios mediante la construcción de carteras de cobertura adecuadas.

Una cartera de cobertura climática busca, en esencia, sobreponderar activos que no estén expuestos al riesgo climático y subponderar aquellos que sí lo estén. Si el riesgo climático aumenta, esta cartera debería apreciarse en el presente; si el riesgo disminuye, debería rendir por debajo del mercado. Si el riesgo no cambia, se trataría de una cartera que reduce el riesgo agregado y, por lo tanto, debería tener una prima de riesgo negativa, similar a un contrato de seguro.

1 Si tomamos como fecha de referencia enero de 2026, podemos establecer que Tesla ha tenido una caída interanual de aproximadamente un 10 % mientras NVIDIA ha experimentado una subida algo mayor al 50 %.

Cuadro 1

Medidas de desempeño sectorial (24 de enero de 2024)

Nombre	Cap. Mercado	PER (P/E)	PER Adel.	PEG	P/Ventas	P/Valor Libros (P/B)	P/Flujo Caja	P/FCF	BPA 5A (%)	BPA 5A Fut. (%)	Ventas 5A (%)	Variación (%)	Volumen
Energía	3350.22B	8.38	9.84	7.07	0.93	1.74	9.45	8.02	26,11	1,19	21,15	1,40	425.09M
Financiero	9674.51B	13.56	12.97	1.38	1.86	1.66	11.67	10.31	9,84	9,81	13,23	0,40	728.43M
Materiales básicos	1934.20B	16.43	14.14	3.28	1.58	2.05	11.23	19.00	16,92	5,01	13,52	0,43	297.17M
Servicios públicos	1234.17B	17.23	15.11	2.46	1.76	1.66	18.91	71.51	3,43	7,00	8,75	1,46	111.37M
Consumo defensivo	3616.44B	21.47	18.32	2.64	1.25	4.18	19.67	23.05	8,27	8,14	6,74	0,76	206.34M
Consumo cíclico	6818.30B	23.11	18.79	1.98	1.53	4.26	8.74	29.19	16,21	11,66	36,34	-0,82	1.03B
Industriales	5137.29B	23.25	18.11	2.16	1.87	4.25	15.74	25.05	13,22	10,75	7,89	0,68	765.69M
Servicios de comunicación	7153.56B	26.19	19.20	1.27	3.24	3.92	12.66	20.21	21,13	20,62	361,04	1,35	515.80M
Salud	7600.43B	32.12	19.10	3.55	1.96	4.30	13.78	23.26	16,25	9,06	78,79	-0,29	1.12B
Inmobiliario	1473.90B	37.07	30.28	4.85	4.18	2.14	18.66	18.17	10,18	7,65	48,79	0,92	201.64M
Tecnología	16870.25B	38.62	25.51	3.10	5.77	8.68	19.74	36.36	18,41	12,45	16,90	0,25	1.35B
Carbón térmico	11.04B	4.31	6.62	—	0.94	1.42	6.75	3.22	29,15	-4,75	11,90	-1,20	3.60M
Refinación y comercialización de petróleo y gas	225.80B	6.41	10.83	—	0.32	1.90	5.40	5.43	25,37	-8,84	22,29	2,07	23.49M
Petróleo y gas integrado	1525.16B	7.52	8.44	—	0.80	1.49	6.53	6.66	29,76	-6,15	16,69	2,12	85.92M
Exploración y producción (E&P)	730.00B	8.13	10.73	1.07	2.17	1.91	22.76	10.77	36,51	7,57	39,10	1,76	155.41M
Transporte y almacenamiento (Midstream)	578.60B	11.33	12.53	1.98	1.31	2.24	28.77	9.89	9,00	5,71	17,45	0,90	85.18M
Perforación de petróleo y gas	27.32B	12.57	12.29	0.35	1.42	0.96	8.76	19.38	36,45	35,68	6,44	0,69	25.75M
Equipos y servicios de petróleo y gas	229.72B	14.58	11.43	0.62	1.40	2.42	11.94	14.60	17,87	23,49	5,92	1,10	70.95M
Uranio	31.89B	81.31	49.03	1.69	14.43	4.56	11.37	99.74	5,09	48,13	12,15	-3,13	37.57M

¿Han olvidado los mercados de capitales la sostenibilidad?

2. RIESGO DE TERMINACIÓN

A continuación, quiero analizar cómo las empresas y, en última instancia, otros agentes económicos deberían gestionar el riesgo de largo plazo. Introduzco esta cuestión mediante un ejemplo. Supongamos que usted es propietario de un hotel de lujo situado en primera línea de playa, como los que se observan en La Jolla, pero también en Miami, Mozambique o las Maldivas. Estos hoteles de lujo frente al mar están expuestos a numerosos riesgos de largo plazo, entre los cuales destaca el aumento del nivel del mar. Si dicho aumento se materializa, el hotel acabará siendo destruido. Por tanto, la vida útil de este activo es incierta. A este tipo de riesgo lo denominamos *riesgo físico*.

Es probable que este hotel se convierta en lo que se denomina un *activo varado* (*stranded asset*), es decir, un activo cuyo valor se reduce a cero en algún momento incierto del futuro.

¿Cuáles son las implicaciones de este hecho? Una consecuencia bastante directa es que, si se estuviera considerando una reforma o modernización del hotel, es posible que no se lleve a cabo si el período de recuperación de la inversión es largo en relación con la esperanza de vida del activo. En consecuencia, cabría esperar que el propietario reduzca el gasto en inversión y mantenimiento. Como resultado, los mismos ingresos podrían generar mayores beneficios debido a la reducción de los costes. De hecho, podría observarse una mejora temporal de los flujos de caja del hotel.

Además, si existen otros hoteles en la misma playa y todos adoptan un comportamiento similar, la oferta agregada se contrae, lo que previsiblemente incrementa los precios. En este contexto emerge cierto poder de mercado, generando un aumento temporal de los beneficios.

¿Estarían los inversores interesados en poseer acciones de este hotel? ¿Cuál sería su valor? Aunque el hotel será destruido en algún momento, genera flujos de caja en el presente, por lo que tiene un valor positivo, determinado como el valor presente descontado de los flujos de caja futuros esperados, de acuerdo con la teoría financiera estándar. Dado que tanto la fecha de terminación como los flujos de caja son inciertos, es razonable suponer la existencia de una prima de riesgo asociada a la tenencia de este activo. No obstante, cabe esperar que tenga cierto valor, aun sabiendo que, en última instancia, su valor convergerá a cero.

A este tipo de riesgo lo denomino *riesgo de terminación*. La idea central es que existe un punto final incierto: no se sabe cuándo ocurrirá ni siquiera si llegará a producirse. Sin embargo, si existe una probabilidad no despreciable de que este escenario se materialice y marque el final de la actividad empresarial, es necesario planificar teniendo en cuenta dicho riesgo.

Existen otras características relevantes del riesgo de terminación que conviene destacar. Si la demanda de vacaciones en la playa aumenta, cabría esperar un incremento de los ingresos, los beneficios y el precio de las acciones del hotel. No obstante, también podría producirse una respuesta de la oferta. Sin embargo, si la terminación es inminente, la respuesta de la oferta será necesariamente limitada.

Dos fuentes fundamentales de fluctuaciones en los precios de las acciones son los cambios en la demanda y los cambios en la fecha esperada de terminación. En consecuencia, los precios de las acciones deberían aumentar cuando la demanda se incrementa o cuando la fecha de terminación se desplaza más hacia el futuro.

Otra posible estrategia para el propietario del hotel sería la diversificación. Por ejemplo, podría considerar invertir los flujos de caja obtenidos en la creación de una planta de vehículos eléctricos. Sin embargo, salvo que el propietario posea competencias específicas en la fabricación de vehículos eléctricos, los inversores probablemente reaccionarían de forma negativa a este tipo de inversión, ya que preferirían seleccionar por sí mismos la empresa de vehículos eléctricos de su preferencia. En este caso, la diversificación daría lugar a una reducción del precio de las acciones y del valor de la empresa combinada. Este fenómeno fue observado de manera particularmente clara en la industria del tabaco, donde se intentaron estrategias de diversificación excesivas que finalmente fueron abandonadas al comprobarse que el valor de las partes superaba al valor del conjunto.

Por último, si el número de empresas u hoteles es reducido, la contracción de la oferta que genera rentas de monopolio no necesariamente se materializará. Existe competencia entre los propietarios: si uno reduce los precios para aumentar su cuota de mercado, los demás obtendrán menores rentas monopolísticas. Esta situación genera presiones hacia la consolidación. Pueden surgir economías de escala, y una gestión coordinada puede facilitar la administración del proceso de declive y la extracción de rentas monopolísticas.

Describo el riesgo de terminación porque considero que presenta una notable similitud con el sector de la energía fósil, que constituye la conexión que deseo establecer. A lo largo del tiempo, numerosas actividades económicas han llegado a su fin. Ejemplos de ello incluyen Kodak, Blockbuster, RCA Victor y la industria del tabaco. Existen múltiples razones para la terminación de estas actividades, entre ellas los avances tecnológicos, la competencia y las políticas públicas.

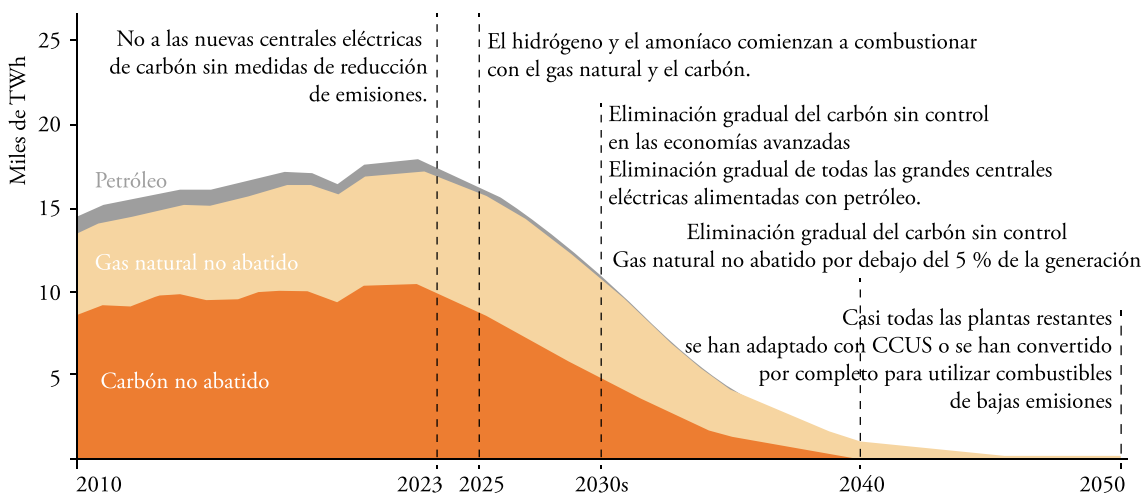
En el contexto energético, a menudo se asume que las empresas de energía fósil se transformarán en empresas verdes. Sin embargo, no considero que este proceso sea sencillo. Su actividad principal consiste en perforar el subsuelo y operar redes de oleoductos y gasoductos, lo cual difiere sustancialmente del desarrollo de parques eólicos, instalaciones solares o infraestructuras

¿Han olvidado los mercados de capitales la sostenibilidad?

eléctricas conectadas a la red. La electricidad fluye a través de redes eléctricas, no de tuberías. No existe una razón clara para suponer que las empresas de energía fósil posean una ventaja comparativa en energías renovables. Es posible que creen filiales dedicadas a actividades verdes

Figura 1

Producción de electricidad a partir de combustibles fósiles no abatidos



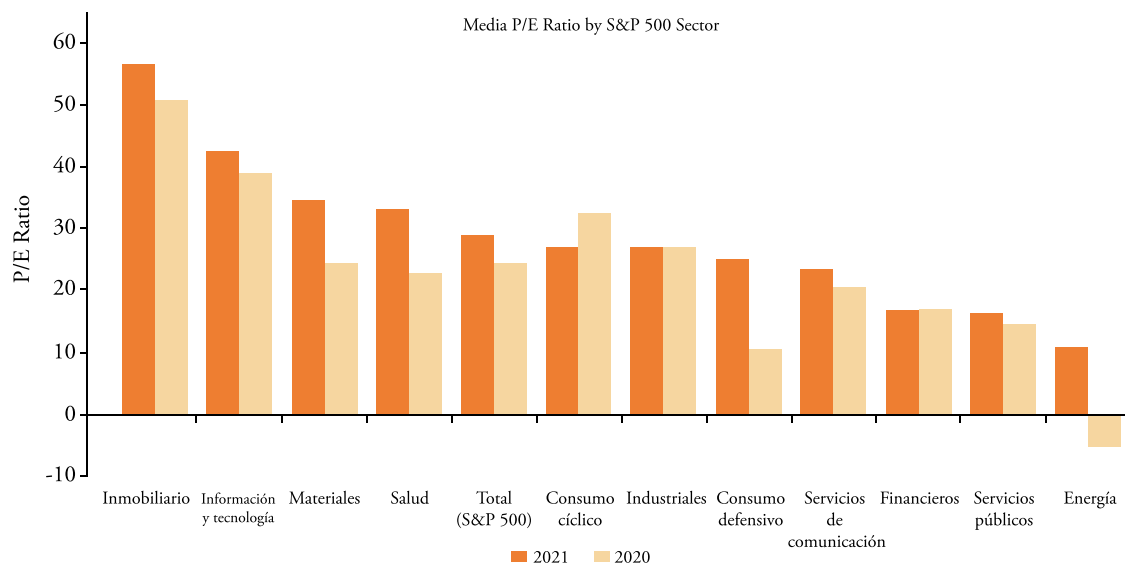
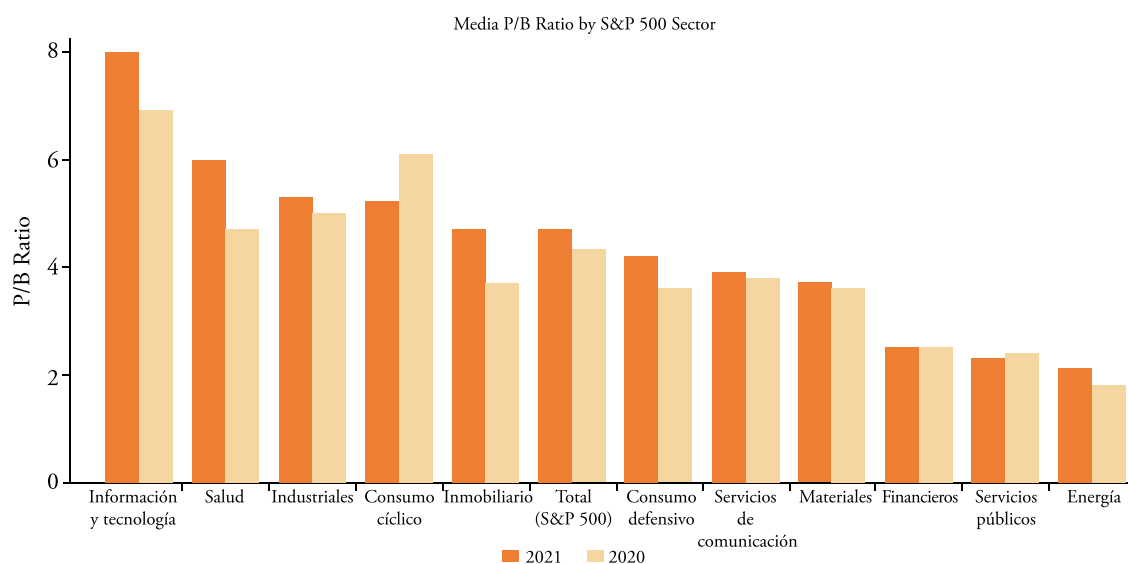
Nota: Producción de electricidad a partir de combustibles fósiles no abatidos baja en un 40 % en 2030 y prácticamente desaparece en 2050 ante la disminución de la actividad de las plantas ya retiradas o alimentadas con Captura y Utilización y almacenamiento del Carbono (*Carbon Capture Utilization and Storage, CCUS*) transformadas para producción de combustibles bajos en carbono.

y posteriormente las escindan, algo que ya se ha observado y que podría resultar una estrategia lógica. Sin embargo, mientras los inversores no estén interesados en invertir en energía fósil cuando es intensiva en carbono, es probable que también se muestren reticentes a invertir cuando dicha energía es “marrón con un ligero matiz verde”.

Finalmente, el riesgo de terminación no solo afecta a empresas, sino también a países cuya principal fuente de ingresos enfrenta un declive estructural. En última instancia, la humanidad en su conjunto enfrenta un riesgo de terminación asociado al cambio climático. A diferencia de empresas o países, no es posible diversificarse frente a este riesgo; la única estrategia viable es reducir su probabilidad mediante cooperación internacional y políticas creíbles que permitan alcanzar emisiones netas cero.

Entonces, ¿es realmente el riesgo de terminación una explicación plausible para las empresas de energía fósil? Consideremos el modelo de la Agencia Internacional de la Energía sobre cómo

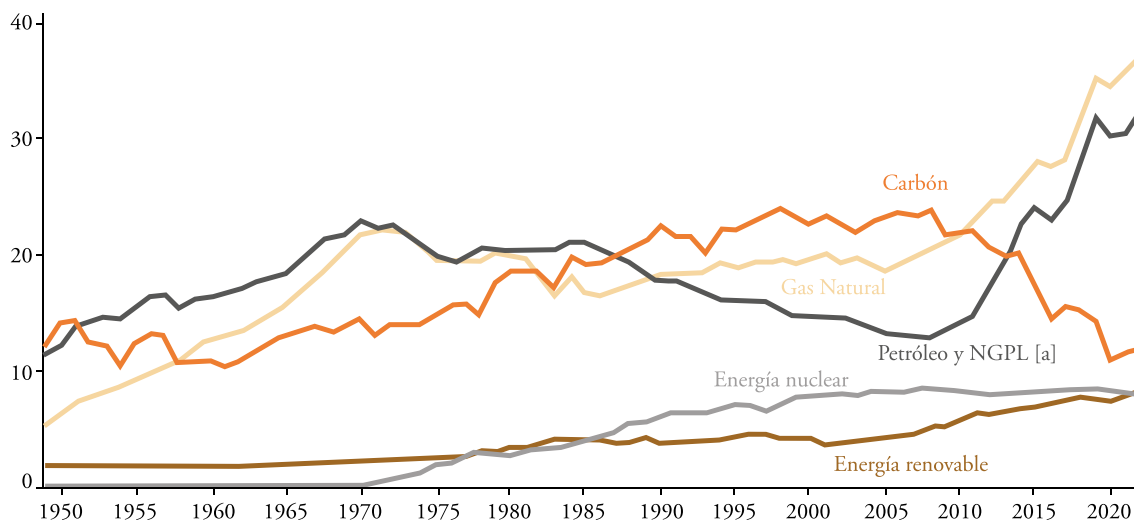
Figura 2

Ratio de valoración por sector S&P 500 durante 2020 y 2021**a) Ratios precio beneficio (P/E)****b) Ratios precio valor (P/B)**

debería ser la transición desde el presente hasta 2050. Existen numerosos pasos intermedios en este proceso, pero la representación mostrada en la figura 1, que incluye carbón, gas y

Figura 3

Evolución del consumo de combustibles fósiles por fuente, 1949-2022



Nota: Producción de electricidad a partir de combustibles fósiles no abatidos baja en un 40 % en 2030 y prácticamente desaparece en 2050 ante la disminución de la actividad de las plantas ya retiradas o alimentadas con Captura y Utilización y almacenamiento del Carbono (*Carbon Capture Utilization and Storage, CCUS*) transformadas para producción de combustibles bajos en carbono.

petróleo utilizados para la generación eléctrica, resulta especialmente ilustrativa. Actualmente nos encontramos en el punto máximo de utilización de estos combustibles, y todo indica que su uso irá disminuyendo progresivamente.

No obstante, cuando se observan gráficos como este, que sugieren que estamos en un punto de inflexión, surge la duda de si realmente creemos que esta transición se materializará. Por otro lado, si tenemos en cuenta los compromisos asumidos para alcanzar emisiones netas cero en 2050, y si existe algún grado de credibilidad en dichos compromisos, entonces existe un riesgo real de que estas empresas energéticas dejen de operar en un horizonte de aproximadamente 25 años. La contrapartida de este proceso es que será necesario movilizar una gran cantidad de capital para financiar las energías renovables, y no está claro que dicho capital esté siendo efectivamente canalizado.

Una vez establecida la plausibilidad económica del riesgo de terminación, pasamos a analizar la evidencia empírica.

¿Existe evidencia empírica que respalde esta noción de riesgo de terminación? Como se ha señalado, los precios de las acciones deberían reflejar el valor descontado de los beneficios

futuros esperados. En consecuencia, las ratios precio-beneficio deberían ser bajos si quedan pocos años de beneficios por delante. Cuando el riesgo de terminación se intensifica, cabría esperar observar ratios P/E reducidos y también ratios precio-valor contable bajos, ya que muchos de los activos de las empresas de energía fósil probablemente se convertirán en activos varados y no estarán plenamente ajustados en los balances contables en comparación con su valor de mercado.

¿Qué se observa cuando analizamos las ratios de beneficios en los años 2020 y 2021? En la figura 2 se aprecia que los sectores energéticos se sitúan sistemáticamente en los valores más bajos en ambos años. De hecho, en 2020 —el año de la pandemia— presentan beneficios negativos. Si se analizan las ratios precio-valor contable, se observa el mismo patrón. Por tanto, estos dos años son coherentes con la hipótesis del riesgo de terminación.

Más recientemente, al observar los datos del 24 de enero de 2024, se identifica nuevamente el mismo fenómeno, tal y como se puede observar en el cuadro 1. Los sectores están ordenados según su ratio precio-beneficio, que corresponde a la segunda columna. De nuevo, el sector energético presenta la ratio P/E más bajo, alrededor de 8, frente a valores cercanos a 38 en el sector tecnológico. En el círculo verde se indica que los beneficios por acción del sector energético crecieron un 26 % en los últimos cinco años, pero se espera que aumenten únicamente un 1 % en los próximos cinco años.

Dentro del sector energético, el carbón térmico presenta la ratio precio-valor contable más bajo, mientras que el equipamiento de petróleo y gas muestra el más alto, con la excepción del uranio, que presenta una ratio P/B extremadamente elevado. No es completamente claro el origen de este comportamiento, pero podría indicar que el uranio enfrenta un horizonte de terminación más lejano.

¿Qué está ocurriendo con los datos de consumo energético? La figura 3 muestra setenta años de datos sobre consumo de energía. Se observa que las energías renovables aumentan de forma sostenida, mientras que la energía nuclear permanece relativamente estable. El consumo de carbón disminuye, lo que sugiere claramente una trayectoria de terminación. Sin embargo, el petróleo y el gas no muestran aún una tendencia descendente clara, lo que plantea la cuestión de si los agentes económicos realmente anticipan una reducción futura de su uso.

Aunque podamos mostrarnos escépticos respecto al cumplimiento pleno de los compromisos de emisiones netas cero para 2050, sigue existiendo un riesgo no despreciable de que la energía fósil, en su conjunto, se enfrente a un proceso de terminación en los próximos 25 años.

Para evaluar este riesgo, podemos considerar las políticas disponibles para llevar a los países y a la economía mundial hacia el objetivo de emisiones netas cero. El Acuerdo de París permitió

¿Han olvidado los mercados de capitales la sostenibilidad?

que cada país decidiera de forma autónoma cómo gestionar este proceso. Distintos países han optado por soluciones diferentes. ¿Cuáles son los tipos de soluciones disponibles? Desde el punto de vista económico, la primera opción es gravar las emisiones de carbono mediante impuestos. La segunda opción es subvencionar las energías renovables. Una tercera opción es la “mandato y control”, un mecanismo que se utiliza ampliamente para controlar las emisiones, los vehículos y las empresas eléctricas, entre otros sectores.

En economías más planificadas, la regulación mediante estándares o mandatos puede ser un instrumento aún más eficaz para reducir las emisiones. La última opción la denomino “esperanza”. También podría llamarse “no hacer nada”. Esta política confía en que el sector privado esté suficientemente motivado para adoptar comportamientos más sostenibles, y en que consumidores, trabajadores, empresas e inversores adopten voluntariamente prácticas más verdes. Los consumidores comprarían productos más sostenibles, los trabajadores preferirían emplearse en empresas más verdes, y las empresas e inversores adoptarían de forma voluntaria objetivos de descarbonización.

Es evidente que parte de este comportamiento ya se está observando. Sin embargo, desde el punto de vista económico, esta estrategia exige demasiado, dado el problema de los comportamientos oportunistas o *free riders*, que probablemente debilitarán este esfuerzo colectivo. La “esperanza” carece de una definición clara sobre cuánto es suficiente. Podemos observar si la dirección del cambio ayuda a mitigar el cambio climático, pero no si será suficiente para alcanzar los objetivos climáticos. Se está pidiendo mucho al sector privado para que actúe voluntariamente, especialmente si estas acciones perjudican la rentabilidad.

Dado que todos estos agentes son previsores, basarán sus decisiones no solo en las políticas actuales, sino también en aquellas que probablemente se implementarán en el futuro. Si los desastres climáticos continúan produciéndose, es posible que el electorado elija políticos con políticas climáticas más ambiciosas. Si estas políticas se retrasan hasta que la urgencia sea extrema, es probable que resulten mucho más costosas que si se implementaran de forma anticipada y planificada.

3. CARTERAS DE COBERTURA FRENTE AL RIESGO CLIMÁTICO

¿Cuáles son las implicaciones para los inversores? Supongamos que los inversores desean cubrirse frente al riesgo climático. Es posible que no quieran estar tan expuestos a este riesgo como lo estarían si mantuvieran la cartera de mercado. ¿Cómo se construyen este tipo de carteras? Existen dos estrategias generales.

La primera consiste en clasificar los activos en grupos con baja exposición al riesgo climático y grupos con alta exposición. Esta clasificación puede basarse en múltiples características, como datos contables, sectoriales o geográficos, que son bien conocidos por los mercados financieros. También podrían utilizarse datos ESG (ambientales, sociales y de gobernanza). Sin embargo, este conjunto de datos es controvertido, ya que no siempre está claro qué se desea medir; suele ser información voluntaria y frecuentemente interpolada por proveedores para cubrir un universo amplio de activos. Aunque se están realizando esfuerzos para mejorar la divulgación, incluso con información obligatoria, los datos actuales no son suficientes para clasificar con precisión los activos.

La segunda estrategia es de naturaleza estadística. Ante la evidencia de un aumento del riesgo climático, se analiza qué acciones suben y cuáles bajan. Este enfoque refleja la visión del mercado sobre el grado de exposición climática de las empresas, que puede diferir de la información ESG o de los datos de emisiones. Una cartera de cobertura debería sobreponderar los activos que se benefician de noticias adversas sobre el clima y subponderar aquellos que se ven perjudicados.

En el Volatility and Risk Institute (VRI) hemos desarrollado una serie de carteras de cobertura climática, cuyo rendimiento publicamos diariamente en nuestra plataforma VLAB. Nos centramos en una cartera denominada *Stranded Asset Portfolio* (SA) o cartera de activos varados. Esta fue propuesta por Bob Litterman. Esta cartera mantiene una posición larga en el ETF del S&P 500 (SPY) y una posición corta compuesta por un 70 % en un ETF de carbón (KOL) y un 30 % en el ETF energético amplio (XLE). Esta cartera se revaloriza cuando las acciones de energía fósil caen y se basa exclusivamente en la clasificación sectorial de las empresas.

Una alternativa más estadística es la cartera de factores eficiente frente al clima (*Climate Efficient Portfolio*, CEP), que se construye como una cartera de fondos sostenibles a partir de nuestra base de datos, la cual contiene unos 200 de estos fondos. El objetivo es obtener una cartera *long-only*, de mínima varianza y máxima correlación con noticias climáticas, condicionada a factores de riesgo estándar, rendimientos del petróleo y la cartera de activos varados. Estas carteras se reequilibran mensualmente y son plenamente invertibles.

Examinamos su comportamiento en el largo y corto plazo e ilustramos los resultados en los cuadros 2 y 3, respectivamente. El cuadro 2 muestra los resultados para SPY, ACWI, SA, CEP y una cartera de cobertura que no diferencia al carbón. Todas las rentabilidades reportadas son fuera de la muestra (*out of sample*). El análisis del rendimiento a largo plazo refleja que estas carteras presentan resultados comparables a los de las carteras de mercado amplio. El rendimiento relativo depende de la intensidad de las noticias climáticas a lo largo del tiempo. En periodos prolongados de aumento del riesgo climático, estas carteras pueden superar al mercado. En particular, la cartera de activos varados presenta la mayor ratio de Sharpe, seguida por SPDR y CEP.

Cuadro 2

Desempeño a largo plazo de las carteras de referencia

Referencia	Horizonte (años)	Rentabilidad (%)	Volatilidad (%)	Ratio Sha
iShares MSCI ACWI ETF	15.98	7.06	20.81	0.13
SPDR S&P 500 ETF Trust	31.13	10.26	18.74	0.31
Stranded Assets	25.25	7.74	22.56	0.34
SPY:US–XLE:US	25.25	–5,01	21.75	–0,23
Climate Efficient Factor Mimicking Portfolio	22.72	10.38	24.48	0.24

En horizontes temporales más cortos (ver cuadro 3), el rendimiento es más variable. Al analizar los rendimientos anuales con cierre comprendido entre el 24/3/24 y el 24/3/23, se observan cambios significativos.

En 2022, por ejemplo, las carteras de cobertura climática tuvieron un rendimiento significativamente peor que el mercado, con mayor volatilidad. Este fue un año marcado por el aumento de los tipos de interés y la lucha contra la inflación, así como por la recuperación posterior a la pandemia y la crisis energética en Europa. En cambio, en 2023 y comienzos de 2024 estas carteras mostraron una recuperación significativa.

Cuadro 3

Desempeño a un año de las carteras de referencia

Referencia	Rentabilidad (%)	Volatilidad (%)	Ratio de Sharpe
iShares MSCI ACWI ETF	–8,07	22.67	–0,56
SPDR S&P 500 ETF Trust	–8,18	24.18	–0,53
Stranded Assets	–25,63	27.08	–0,95
SPY:US–XLE:US	–40,64	32.24	–1,26
Climate Efficient Factor Mimicking Portfolio	–23,75	37.99	–0,75
iShares MSCI ACWI ETF	27.73	11.36	2.00
SPDR S&P 500 ETF Trust	34.44	11.53	2.55
Stranded Assets	20.04	16.52	1.21
SPY:US–XLE:US	7.13	19.75	0.36
Climate Efficient Factor Mimicking Portfolio	14.27	13.57	0.68

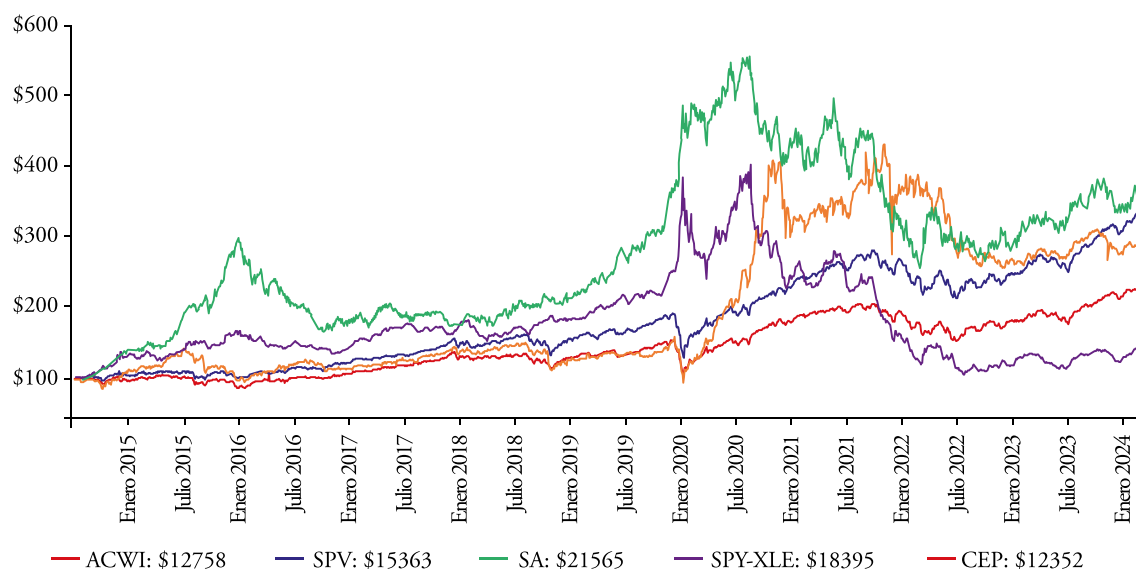
Durante la pandemia, el riesgo de terminación parecía particularmente elevado, como lo demuestra el colapso de la demanda energética y los precios negativos del petróleo. A medida que la pandemia remitió, los precios del petróleo aumentaron y el riesgo de terminación se redujo, impulsando

aún más las valoraciones bursátiles. El bajo rendimiento relativo de los fondos sostenibles en 2022 se explica fundamentalmente por su infraponderación en energía fósil, un sector que obtuvo resultados excepcionalmente buenos. La pregunta clave es por qué el petróleo y el gas tuvieron tan buen desempeño. La hipótesis es que esto responde al riesgo de terminación: las empresas reducen la inversión, aumentan los flujos de caja y, como consecuencia, los precios de las acciones suben. Además, los *shocks* de demanda y oferta refuerzan este efecto. Concretamente, durante entre 2021 y 2022 hubo una escalada en el precio y la volatilidad del gas, la electricidad y el petróleo en Europa a raíz de la invasión a Ucrania por parte de Rusia y la utilización del gas como herramienta de presión.

La figura 4 muestra la evolución de las distintas series durante el periodo comprendido entre 2014 y 2024. La cartera de activos varados, representada en verde, tuvo un desempeño extremadamente bueno antes y durante la pandemia, aumentando de forma pronunciada desde enero de 2019 hasta finales de 2020. Posteriormente descendió hasta enero de 2023, pero desde entonces se ha recuperado. El CEP, representado en naranja, aumentó con fuerza desde marzo de 2020 hasta enero de 2022, cayó en 2022 y se ha recuperado en 2023.

Figura 4

Evolución del beneficio proporcionado por las distintas carteras en la última década



3.1. Desempeño de los fondos sostenibles

¿Por qué tuvieron malos resultados en 2022, pero un buen desempeño en 2020 y en 2023/4? Hemos analizado dos carteras de referencia de cobertura climática que muestran patrones

¿Han olvidado los mercados de capitales la sostenibilidad?

similares, aunque con diferencias importantes. Si observamos los más de 200 fondos sostenibles durante el mismo período de la muestra en VLAB, las correlaciones son bastante altas, pero existe una divergencia significativa entre los fondos. El promedio equiponderado de todos los fondos presenta sistemáticamente un rendimiento inferior al del SPY y comparte muchas de sus mismas características. El promedio muestra un sólido desempeño durante la pandemia, seguido de un rendimiento débil en 2022, que posteriormente se fortalece en 2023 y en la primera mitad de 2024. Dentro del conjunto de fondos sostenibles (figura 4), hay muchos con un desempeño muy superior, al menos durante partes del período.

La razón por la que los fondos sostenibles han tenido un rendimiento inferior es simplemente que están infraponderados en energía fósil, y el sector del petróleo y gas tuvo un mejor desempeño. La pregunta es: ¿por qué el petróleo y el gas superaron al mercado? La hipótesis es que se trata de una respuesta natural al riesgo de terminación. El riesgo de terminación incentivaría a las empresas a reducir su inversión y, por lo tanto, a aumentar el flujo de caja. Y si el flujo de caja aumenta de manera significativa, los precios de las acciones también suben. Además, si se produce un *shock* de demanda, esto también impulsa nuevamente estos precios. Asimismo, es muy probable que el riesgo de terminación se percibiera como más severo en 2020, durante la pandemia, que en los años posteriores. Recordemos que durante la pandemia dejamos de conducir, no nos desplazábamos al trabajo o a la escuela, volamos mucho menos y llevamos el precio del petróleo a niveles tan bajos que, en un momento dado, los futuros del petróleo llegaron a ser negativos.

En este contexto, el riesgo de terminación debe haber sido altamente plausible y costoso.

Así, a medida que la pandemia fue remitiendo, los precios del petróleo aumentaron y el riesgo de terminación presumiblemente retrocedió, lo que condujo a precios de las acciones aún más altos.

Este efecto se ve amplificado por el hecho de que no contamos con un impuesto integral sobre las emisiones. A medida que los precios y las ganancias aumentaron, las recompensas fueron a parar a los accionistas en lugar de a los recaudadores de impuestos. Un impuesto al carbono habría reducido la rentabilidad de la energía fósil y proporcionado recursos para otras políticas públicas o para reducciones impositivas.

Los países también enfrentan riesgo de terminación. El riesgo para un país es que su industria principal llegue a extinguirse. Si pensamos en Oriente Medio, vemos que se enfrentan a este problema a diario. Están tratando de utilizar sus enormes flujos de caja para diversificar sus economías. Esto se observa en la inversión en turismo, deportes y viajes, educación, artes y una amplia variedad de sectores en los que están invirtiendo fuertemente, de modo que cuando desaparezcan los ingresos energéticos, cuenten con actividades económicas alternativas. Además, la OPEP ha reducido la producción en los últimos años, tal como se predecía a partir del riesgo de terminación.

¿Y qué ocurre con Rusia? Rusia también ha reducido el suministro de energía fósil. Cortó el suministro de gas a Europa, incluido el gasoducto que atraviesa Ucrania, redujo las reservas en Europa y las sanciones lo redujeron aún más. Decidieron, en mi opinión, que para diversificar su economía la mejor opción era invadir Ucrania, ya que Ucrania les proporcionaría mayores capacidades de exportación y recursos naturales. Creo que cometieron un error de cálculo respecto a lo fácil o difícil que sería esto. No obstante, considero que la cuestión del riesgo de terminación en Rusia es muy seria: dentro de una década, Rusia no será financieramente tan fuerte como lo era hace dos años cuando inició este proceso, y Europa contará con mucha más energía verde y, por lo tanto, será menos vulnerable a las restricciones del gas ruso. Esta no es la única explicación de la invasión, pero creo que forma parte de los incentivos. Este argumento sugiere que el cambio climático está aumentando la probabilidad de conflictos bélicos.

Por último, cabe mencionar a Irán, que también enfrenta un riesgo de terminación en su industria principal. Dispone de elevados flujos de caja, pero los está utilizando para desestabilizar Oriente Medio. No está claro cómo se espera que esto contribuya a diversificar la economía iraní, pero, nuevamente, el cambio climático está fomentando los conflictos armados.

Como defensores del clima, ¿deberíamos sentirnos desalentados por lo que está ocurriendo? Aunque nuestras inversiones sostenibles tuvieron un mal desempeño durante 2021-2022, los precios de la energía fueron elevados y, desde el punto de vista de la mitigación del cambio climático, los altos precios de la energía fósil son positivos. Los precios elevados de la energía desincentivan la compra de bienes producidos con altas emisiones. No creo que a ningún economista le sorprenda que una mitigación climática eficaz incluya el aumento del precio de los bienes con altas emisiones. Cuando las personas dicen: “vas a la gasolinera con tu SUV y es carísimo, ¿qué se supone que debo hacer?”, esa es precisamente la respuesta: se supone que se debe conducir menos, comprar un vehículo eléctrico o, al menos, un automóvil eficiente en el consumo de combustible. Sabemos que este impacto recae de manera desproporcionada sobre las personas de menores ingresos. Ese es un problema que debe abordarse, pero es una cuestión distinta. Por lo tanto, los altos precios de la energía no constituyen un problema para la mitigación climática.

¿Significa esto que ya no deberíamos invertir en estas carteras de cobertura? Suponiendo que sigamos intentando cubrir el riesgo climático, los cambios en el riesgo de terminación moverán una cartera de cobertura en la dirección correcta. Si el riesgo de terminación aumenta y la fecha se acerca, las acciones de petróleo y gas caerán y tenderán a cero. Eso es exactamente lo que se desea que ocurra en la cartera, por lo que se debería seguir manteniendo posiciones cortas en empresas expuestas al riesgo de terminación.

Observamos que Putin ha acelerado la descarbonización, el proyecto de ley climática de Biden también lo hizo, y el Pacto Verde de la Unión Europea también tiene este objetivo. Estos

¿Han olvidado los mercados de capitales la sostenibilidad?

acontecimientos geopolíticos realmente nos movieron en una dirección favorable al clima. Está por ver qué sucederá con los cambios promovidos por Trump o los nuevos balances políticos en Europa.

También podríamos preguntarnos por el riesgo de terminación de la especie humana. Tenemos una estrategia para una empresa que enfrenta la terminación y una estrategia para un país que enfrenta la terminación, pero ¿cuál es la estrategia para la humanidad? No podemos diversificarnos. No podemos reducir nuestra inversión. Esas no son buenas opciones. Lo único que podemos reducir es la probabilidad. Necesitamos reducir la probabilidad de la terminación.

El Acuerdo de París proporciona una hoja de ruta. La ciencia afirma que, si el mundo puede alcanzar emisiones netas cero para 2050, evitaremos las peores consecuencias del cambio climático. No podemos hacer cumplir estos compromisos salvo mediante la diplomacia y la opinión pública. Por lo tanto, la mejor manera de reducir el riesgo de terminación es que los principales emisores colaboren, cooperen y trabajen juntos.

Si los cuatro mayores emisores logran alcanzar las emisiones netas cero, los países más pequeños los seguirán. Tendrán acceso a nuevas tecnologías y una fuerte presión moral. Con suerte, si existen países que actúan como *free riders*, serán pocos y pequeños. Los mayores emisores son China, Estados Unidos, Europa e India. Europa está trabajando activamente en la descarbonización, Estados Unidos inició con Biden una política climática, China ha invertido fuertemente en tecnología verde, pero aun así enfrenta un rápido aumento de sus emisiones, e India aún tiene mucho trabajo por hacer. Tengo la sensación de que la clave es que Estados Unidos y China trabajen juntos. Hay muchas maneras de hacerlo y el futuro de todos nosotros depende de que esto tenga éxito.

Understanding the Dispersion of Environmental Scores across Industrial Sectors*

Juan Ignacio Peña**, Kurt A. Desender*** and Bing Guo****

Abstract

This paper aims to explore factors that may explain the dispersion of environmental ratings across industrial sectors. The paper studies external dispersion –the discrepancy in E ratings across raters–, and internal dispersion –the discrepancy across E dimensions within a single rater. We classify industrial sectors by their degree of “greenness” or “brownness” based on their unscaled greenhouse gas (GHG) emissions. We study the impact of characteristics such as the ownership structure (including common ownership, institutional ownership, or family ownership), the proportion of independent directors on the board, the degree of CEO duality, and board- and executive-level gender diversity on the dispersion of environmental ratings. External dispersion mainly reflects cross-rater comparability problems driven by firm complexity, ownership structure, and political context, whereas internal dispersion captures persistent within-rater heterogeneity across environmental dimensions. External dispersion primarily signals methodological divergence across providers, whereas internal dispersion reflects informative multidimensional environmental performance.

Keywords: environmental scores, rating discrepancies, sustainability metrics, data providers.

1. INTRODUCTION

The global financial environment has focused on Environmental, Social, and Governance (ESG) factors as key elements in company valuation for several years (Gibson *et al.*, 2022). For example, over the past decade, ESG considerations have become deeply embedded in European financial markets, supported by regulatory initiatives such as the EU Taxonomy, the Sustainable Finance Disclosure Regulation (SFDR), and the Corporate Sustainability Reporting Directive (CSRD). Institutional investors increasingly incorporated ESG ratings into portfolio construction, rising from 12% in 2018 to 43% in 2022 (Wong *et al.*, 2023), while policymakers promoted standardized sustainability disclosure as a cornerstone of capital market integration. However, the policy momentum behind this conventional ESG framework is now being reassessed. Evidence of substantial divergence across ESG rating providers has raised concerns among regulators, supervisors, and market participants about

* We acknowledge research assistance from Jack Louis Christie and financial support from FUNCAS through the ESGPE 2024/00644/001 grant. We thank the Editor, Pedro Linares, for useful comments.

** Universidad Carlos III de Madrid, Department of Business Administration, ypenya@eco.uc3m.es

*** Universidad Carlos III de Madrid, Department of Business Administration, kdesende@emp.uc3m.es

**** Universidad Carlos III de Madrid, Department of Business Administration, bguo@emp.uc3m.es

comparability, transparency, and the risk of misallocation of capital. Studies documenting ESG rating variability, including the present analysis, have contributed to a shift in the debate: from expanding the use of ESG ratings to scrutinizing their methodological foundations and considering stronger harmonization and oversight mechanisms.

Kotsantonis and Serafeim (2019) indicate that the most significant obstacle to using ESG data in investment decisions is the lack of comparability of ESG metrics across companies and over time. Avramov *et al.* (2022) show that demand for equities decreases and market risk increases if there is uncertainty in ESG ratings. Billio *et al.* (2021) find that disagreement among rating scores disperses the effect of ESG preferences on asset prices to the point that ratings no longer affect financial performance. Berg *et al.* (2022) introduce the concept of “aggregate confusion,” referring to the divergence in ESG ratings stemming from differing methodologies and definitions across rating agencies. They document the rating divergence and decompose it into scope, measurement, and weight contributions.

The motivation for this paper stems from the results in Peña *et al.* (2025), which study the external (*i.e.*, disagreement in environmental scores across rating agencies) and internal (*i.e.*, disagreement across environmental dimensions within a single rating provider) dispersion of the E ratings. They document that large firms, firms with more tangible assets, firms with strong expected future performance, firms with credit ratings, and those in Democratic-leaning states experience higher rating discrepancies. In contrast, more profitable firms see lower disagreement, and the covariation between the growth opportunities and disagreement is significant. Moreover, although internal and external discrepancies are related to similar factors, the direction and nature of their association differ, highlighting the complexity of factors driving rating disagreements. However, they do not examine differences across industries in GHG emissions.

While prior research documents substantial disagreement across ESG rating providers and identifies firm-level determinants of rating dispersion, existing studies do not examine whether the drivers and interpretation of rating disagreement differ systematically across sectors with different environmental intensity. In particular, little is known about whether rating dispersion reflects measurement problems or economically meaningful heterogeneity in environmental performance conditional on sectoral emissions exposure. The novelty of this paper lies in classifying industries into two categories, “green” and “brown”, and examining whether the economic determinants and interpretation of environmental rating dispersion differ across these sectoral environments. The criteria for assigning industries to a category depend on an emissions-based indicator. Higher values of this indicator correspond to industries in which firms exhibit higher CO₂-equivalent emissions. We investigate which firm characteristics are associated with the internal and external dispersion of E ratings across both categories, highlighting their differences and similarities.

Disagreement across rating agencies is more pronounced in brown (high-emission) industries, especially due to ownership structure and political factors. In greener industries, cross-agency differences are more influenced by profitability and institutional investors, suggesting that better disclosure helps reduce confusion. Within a single rating agency, however, green industries show more variation across environmental dimensions, reflecting real strengths and weaknesses. In brown industries, internal ratings tend to be more uniform, particularly in firms with strong family control or high credit quality. In short, brown sectors face more comparability problems across raters, while green sectors display more meaningful internal differences in environmental performance.

The main message of this paper is straightforward but important: not all disagreement in environmental ratings means the same thing. When different rating agencies give very different scores to the same company, the gap often comes from differences in methods, definitions, or weighting, but not necessarily from real differences in environmental performance. This type of disagreement can confuse investors and make it harder to compare companies. By contrast, when E ratings produced by a single agency show that a company performs well in some environmental areas but poorly in others, this may reflect real strengths and weaknesses rather than measurement problems.

This distinction matters for decision-makers. For policymakers, the priority should be improving transparency and consistency across rating providers to prevent avoidable methodological differences from distorting markets. For investors, the lesson is clear: a single headline ESG score is not enough. Therefore, understanding why ratings differ is essential for sound capital allocation. For managers, clearer and more structured disclosure can reduce cross-agency disagreement while still revealing the company's environmental challenges. In short, disagreement in ESG ratings is not automatically a flaw. Nevertheless, knowing its source is crucial for better policy, better investment decisions, and better corporate strategy.

The rest of the paper is structured as follows: Section 2 discusses related literature, Section 3 details the database and the criteria for building the green and brown categories. Section 4 discusses the empirical results, and Section 5 concludes.

2. RELATED LITERATURE

Chatterji *et al.* (2016) show considerable disagreement among the ESG ratings of the top data providers. Huber and Comstock (2017) note that providers' methodologies, scopes, and coverage of reports and ratings vary significantly. Christensen *et al.* (2021) present evidence that companies that provide more information about their ESG factors exhibit the most significant disagreement (or variation) in their ESG ratings. In addition, the greater

the disagreement between ESG ratings, the higher the volatility of equity returns and the lower the likelihood of resorting to external financing. Therefore, this indicates that more disclosure of ESG data increases disagreement in ESG ratings rather than decreases it. Notably, disagreement is greatest in the E and S categories, whereas agreement among raters is higher in the G category. One possible reason for these results is that more information increases the likelihood of different interpretations. This issue is important because, without agreement on what constitutes good ESG performance, it is doubtful that ESG ratings will provide relevant information for investment decisions.

In addition, the ESG rating agency industry raises controversial issues, such as potential conflicts of interest and *greenwashing*. Theory and evidence (Pedersen *et al.*, 2021) suggest that the impacts on several firm value measures of separate E, S, and G or ESG may differ. For instance, G ratings correlate positively with equity value, S ratings correlate negatively, and E ratings present mixed results. This paper focuses on the E dimension for several reasons. First, the E-dimension is directly linked to observable factors such as carbon emissions and water use, enabling a more objective and measurable assessment of environmental performance than social or governance performance (Gibson *et al.*, 2021). Second, the growing global emphasis on environmental sustainability, as reflected in international agreements such as the Paris Climate Agreement, has heightened the importance of E-related disclosures and performance metrics. Third, compared to the social dimension, which encompasses a broad and diverse set of factors, such as labor rights, diversity, and community impact, E is a more narrowly defined measure, facilitating better comparability across firms.

3. DATA

In this section, we describe the data sources and the methodology for assigning industries to the “green” and “brown” categories.

3.1. Data Sources

This paper uses a sample of environmental performance data from three databases covering 2002 to 2019 for US-listed companies to extract environmental performance measures. As a novelty relative to the extant literature, the sample includes a database constructed by Sautner *et al.* (2023) that focuses explicitly on the environmental dimension and two commercial providers, LSEG Data & Analytics (previously known as Refinitiv or Asset4) and MSCI ESG Research, whose data have been widely used in many academic studies. We consider all firms for which data is available for all three ratings from 2002 to 2019. The initial sample includes 15,553 firm-year observations from 3,244 firms. Our first data source provides a commercial ESG rating by LSEG Data & Analytics, a financial data company owned by the London

Stock Exchange Group (LSEG). Its environmental performance scores assess firms on various sustainability criteria, including resource use, environmental innovation, and climate-related disclosures. Our second source is another commercial ESG rating provided by the MSCI ESG Research database, initially developed by KLD Research & Analytics, Inc., and later acquired by MSCI (Morgan Stanley Capital International). This rating provider focuses on several ESG dimensions, strengths, and concerns. This database is a long-established, detailed repository of ESG indicators, particularly in the U.S. market. Our focus is on environmental strengths and concerns, including corporate environmental practices such as pollution, renewable energy use, and compliance with environmental regulations. Our third non-commercial data source is the climate change exposure constructed by Sautner *et al.* (2023). To construct this measure, Sautner *et al.* (2023) focus on corporate quarterly earnings calls to assess firms' climate change risks and opportunities. They capture a firm's exposure to a topic by measuring the proportion of the conversation during earnings calls dedicated to that subject (Hassan *et al.*, 2019). Compared to the commercial environmental ratings, this measure provides an alternative perspective by focusing on the attention financial analysts and management devote to climate change topics during conference calls. The external dispersion is computed as the standard deviation of the standardized individual ratings. Table 1 provides further information about ESG data providers, the firms and periods they cover, and the number of firms and rating scales.

Table 1

ESG Data Providers

Data Provider	Country	Rating Scale	Period Covered	Number of Firms	Pillars	Rating Style
LSEG	95 Countries	0-100	2002-2024	11,672	E, S, G, Total	Index
MSCI	141 Countries	0-10	Jan. 1999-Dec. 2024	29,167	E, S, G, Total	Index
Sautner <i>et al.</i> (2023)	88 Countries	[0-0.075] %	2001-2023	15,198	E	Proportion

Note: This Table presents information on ESG data providers. The sample is from 2002 to 2019.

We then incorporate firms' financial and accounting data (*e.g.*, total assets, leverage, capital expenditures, or cash holdings) from Compustat North America to the initial sample. Data on stock price volatility come from CRSP. We also retrieve data on board characteristics, institutional investors' ownership structures, and Moody's credit ratings from LSEG. The final sample includes 9,046 firm-year observations, which correspond to 1,472 firms.

3.2. Green and Brown Industries

Refinitiv provides three variables related to CO₂ emissions: *co2emistorevscore*, *co2equivalentemissions*, and *co2eqemitoevic*. The *co2emistorevscore* is the relative score Refinitiv assigns to this category; *co2equivalentemissions* is the measure of total CO₂-equivalent emissions, defined as the sum of a company's direct and indirect greenhouse gas emissions, measured in tons; and *CO2eqemitoevic* measures CO₂ and equivalent emissions to Enterprise Value Including Cash (EVIC). There are 133 industry codes with CO₂ emissions data.

In this paper, we focus on the variable *co2equivalentemissions* because the objective is to classify industries based on their contribution to aggregate environmental pressure rather than their emissions efficiency relative to firm size. We calculate the median CO₂-equivalent emissions for each industry code using firm-year observations. We define a firm-year observation as having high CO₂-equivalent emissions if its measure is above the median across all firm-year observations in its industry, and low otherwise. The variable *High_co2equivalentemissions* is equal to 1, and *Low_co2equivalentemissions* is equal to 0. Next, we compute a score as follows. For instance, if in one industry there are six (firm-year) observations, three are *high_co2equivalentemissions*, two are *low_co2equivalentemissions*, and one has no information, then the score for this industry would be $(1+1+1+0+0)/5=0.6$ across all six observations. The higher the score, the more firms in that industry have high CO₂-equivalent emissions. In other words, the score captures the prevalence of high-emission firm-years relative to that industry's own distribution. After calculating the score for each industry code, we take the median across all industry scores. One industry is defined as green (brown) if its score is lower (higher) than or equal to the median. Therefore, an industry is deemed green (brown) when it shows a lower (higher) prevalence of high-emission firm-years relative to its within-industry distribution. The indicator variable for green industries using this method is *greenind_co2eqvem*. The classification of green and brown industries using this indicator is in Table 2.

Table 2

Green and Brown Industries

Brown	Green
Aerospace & Defense	Advanced Medical Equipment & Technology
Agricultural Chemicals	Advertising & Marketing
Airlines	Airport Operators & Services
Aluminum	Apparel & Accessories
Appliances, Tools & Housewares	Apparel & Accessories Retailers
Auto & Truck Manufacturers	Auto Vehicles, Parts & Service Retailers
Auto, Truck & Motorcycle Parts	Banks

Table 2 (continuation)

Green and Brown Industries

Brown	Green
Brewers	Biotechnology & Medical Research
Coal	Broadcasting
Commodity Chemicals	Business Support Services
Computer & Electronics Retailers	Business Support Supplies
Computer Hardware	Casinos & Gaming
Construction Materials	Commercial Printing Services
Construction Supplies & Fixtures	Commercial REITs
Consumer Goods Conglomerates	Communications & Networking
Courier, Postal, Air Freight & Land	Construction & Engineering
Department Stores	Consumer Lending
Discount Stores	Consumer Publishing
Diversified Chemicals	Corporate Financial Services
Diversified Industrial Goods Wholesale	Distillers & Wineries
Diversified Mining	Diversified Investment Services
Drug Retailers	Diversified REITs
Electric Utilities	Employment Services
Electrical Components & Equipment	Entertainment Production
Electronic Equipment & Parts	Financial & Commodity Market Operator
Environmental Services & Equipment	Financial Technology (Fintech)
Food Processing	Fishing & Farming
Food Retail & Distribution	Footwear
Gold	Forest & Wood Products
Ground Freight & Logistics	Healthcare Facilities & Services
Home Furnishings	Heavy Electrical Equipment
Home Improvement Products & Services	Heavy Machinery & Vehicles
Hotels, Motels & Cruise Lines	Highways & Rail Tracks
Household Electronics	Home Furnishings Retailers
Independent Power Producers	Homebuilding
Integrated Oil & Gas	Household Products
Integrated Telecommunications Services	IT Services & Consulting
Investment Holding Companies	Industrial Machinery & Equipment
Iron & Steel	Integrated Hardware & Software

Table 2 (continuation)

Green and Brown Industries

Brown	Green
Marine Freight & Logistics	Investment Banking & Brokerage Services
Mining Support Services & Equipment	Investment Management & Fund Operators
Multiline Utilities	Leisure & Recreation
Natural Gas Utilities	Life & Health Insurance
Non-Alcoholic Beverages	Managed Healthcare
Non-Paper Containers & Packaging	Marine Port Services
Oil & Gas Drilling	Medical Equipment, Supplies & Distribution
Oil & Gas Exploration and Production	Miscellaneous Educational Service Providers
Oil & Gas Refining and Marketing	Miscellaneous Specialty Retailers
Oil & Gas Transportation Services	Multiline Insurance & Brokers
Oil Related Services and Equipment	Non-Gold Precious Metals & Minerals
Paper Packaging	Office Equipment
Paper Products	Online Services
Passenger Transportation, Ground & Sea	Personal Services
Personal Products	Professional Information Services
Pharmaceuticals	Property & Casualty Insurance
Phones & Handheld Devices	Real Estate Rental, Development & Operation
Renewable Fuels	Real Estate Services
Restaurants & Bars	Recreational Products
Specialty Chemicals	Reinsurance
Specialty Mining & Metals	Renewable Energy Equipment & Services
Textiles & Leather Goods	Residential REITs
Tires & Rubber Products	Schools, Colleges & Universities
Tobacco	Semiconductor Equipment & Testing
Water & Related Utilities	Semiconductors
Wireless Telecommunications Services	Software
	Specialized REITs
	Toys & Children's Products
	Uranium

Note: This Table presents the classification of industries based on CO₂-equivalent emissions, with total CO₂-equivalent emissions serving as the classification criterion. The sample is from 2002 to 2019.

It is important to note that this classification is based on the prevalence of high emitters within each industry's distribution, not on absolute emissions. Therefore, the classification captures within-industry prevalence of high emitters, not absolute emission levels across industries.

4. RESULTS

This section presents empirical results. We emphasize that both external and internal dispersion of environmental ratings may reflect either informational content or noise, depending on context. The regressions that follow, therefore, document systematic associations rather than causal interpretations of rating disagreement. Ownership structure, governance features, profitability, and disclosure quality are likely jointly determined, and rating disagreement may both influence and respond to firms' reporting practices. For example, institutional investors may select firms with more standardized reporting practices, while firms experiencing greater rating disagreement may adjust disclosure policies or governance structures in response. Similarly, profitability may proxy for investments in internal reporting systems rather than directly influencing rating dispersion. Accordingly, the estimates capture equilibrium correlations in cross-rater disagreement rather than structural effects of firm characteristics on rating dispersion.

4.1. External Dispersion

This section reports regression results for the U.S. 2002-2019 sample, distinguishing between green and brown industries based on unscaled greenhouse gas emissions. The dependent variable is external environmental rating dispersion, capturing disagreement in environmental (E) scores across rating agencies. External dispersion primarily reflects limits to comparability across rating providers. Although such disagreement may reduce transparency from a policy perspective, it may also reflect genuine uncertainty or alternative methodological interpretations of environmental performance. By contrast, internal dispersion need not be interpreted negatively: higher within-rater dispersion may reflect informative multidimensionality in environmental performance rather than measurement error. Accordingly, reducing external dispersion is a natural regulatory objective, whereas internal dispersion can contain economically meaningful information about environmental trade-offs. In the following Table, positive (negative) coefficients indicate higher (lower) dispersion. Statistical significance is denoted by *, **, and *** for the 10%, 5%, and 1% levels, respectively. The regression results are in Table 3. We also report the Economic Impact (EI) of each independent variable in both Table 3 and Figure 1. The EI illustrates the percentage impact on the dependent variable of a typical variation in the explanatory variable¹.

¹ The Economic Impact (EI) metric measures the practical influence one variable has on another. In other words, it illustrates the percentage impact on Y of a typical variation in X. The higher the EI, the more relevant the explanatory variable is in explaining variations in the dependent variable. This metric helps prioritize which variables are most important for business decisions, policy changes, or resource allocation. A variable might be statistically significant but have a low Economic Impact coefficient, or vice versa.

Table 3

Associations between the external dispersion of E ratings

Elements	Green industries	EI (%)	Brown industries	EI (%)
Lagged ext. dispersion	+ ***	0.605	+ ***	0.59
Firm size	+ ***	1.97	+	0.46
Profitability	- ***	-13.39	- *	-13.38
Volatility	+	4.20	-	-4.20
Credit ratings	+	2.12	+ *	4.77
Tobin's Q	+ *	0.69	+	1.00
PPE	-	-7.54	+	8.84
Leverage	- *	-6.16	+	0.04
Board independence	+	0.04	+	0.04
Board size	+	0.48	+	0.33
CEO duality	-	-2.24	-	-1.87
Board gender diversity	+	0.05	+	0.11
Family ownership	-	-5.23	+ *	19.36
Institutional ownership	- **	-10.33	- **	-14.63
Common ownership	-	-9.4	-	-0.88
Allred	+	2.93	-	-0.48
Allblue	+	2.07	+ **	4.47
No. of observations	4,248	4,248	3,424	3,424

Note: This Table presents associations between the external dispersion of E ratings (*i.e.*, dispersion across data providers), measured by the logarithm of the standard deviation of those ratings, and firm characteristics. Robust standard errors are estimated. The signs of the coefficients and their Economic Impacts are reported. EI is the Economic Impact measure. Standard errors are clustered at the firm level. ***, ** and * denote statistical significance at 1%, 5% and 10% level.

■ *Persistence of external rating dispersion*

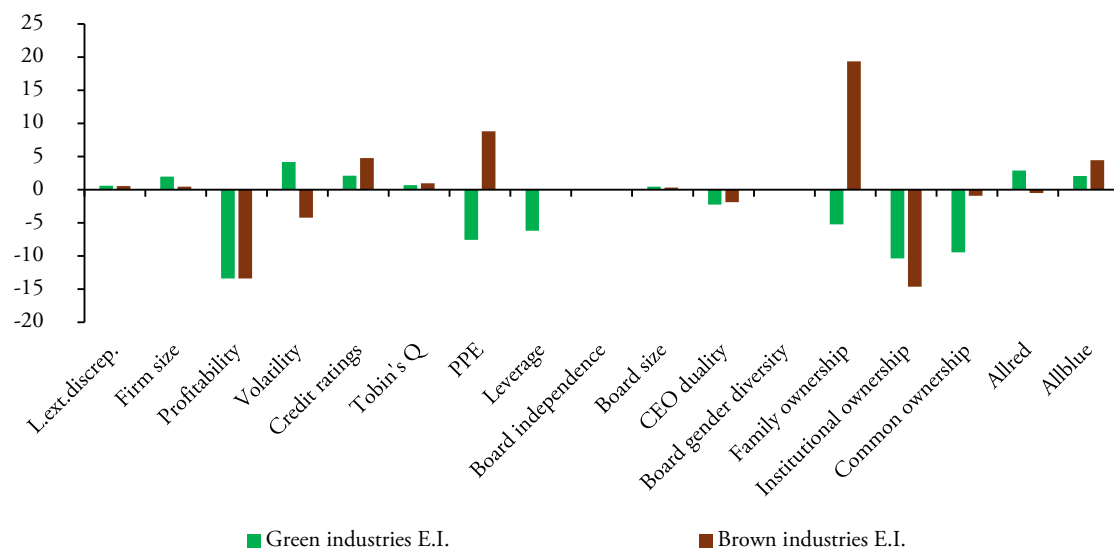
For both green and brown industries, lagged external discrepancy is positive and highly significant at the 1% level. This result indicates strong persistence in cross-rater disagreement across sectors with varying environmental intensity. External dispersion in environmental ratings, therefore, appears to reflect persistent information frictions rather than short-run fluctuations, even within a single regulatory setting. However, the EI is very low.

■ *Firm characteristics*

Firm size is positively and significantly associated with rating dispersion (at the 1% level) in green industries, but not in brown industries. This result suggests that complexity-driven disagreement among raters is particularly pronounced in firms operating in environmentally friendly sectors, where boundaries and materiality may be less clearly defined.

Figure 1

External Dispersion: Economic Impact



Profitability is negatively associated with dispersion in both green (significant at the 1% level) and brown (significant at the 10% level) industries. More profitable firms exhibit lower cross-rater disagreement, consistent with clearer disclosure practices or more stable operational environments that facilitate convergence in environmental assessments. Credit ratings are positively related to external dispersion (significant at the 10% level) in brown industries only. In emissions-intensive sectors, stronger credit quality does not reduce rating disagreement; instead, raters may differ in how they incorporate financial strength alongside environmental impact and transition risk.

Tobin's Q is positively associated with external dispersion (at the 10% level) in green industries, but it is not significant in brown industries. This pattern suggests that growth-oriented firms in cleaner sectors generate more divergent interpretations of forward-looking environmental strategies and commitments. Leverage is negatively related to dispersion in green industries (significant at the 10% level). Greater financial discipline among firms in environmentally friendly sectors may be associated with more standardized reporting or reduced ambiguity in environmental disclosures. Tangibility (PPE) does not present a meaningful impact.

■ Ownership and governance characteristics

Institutional ownership is negatively and significantly associated with external rating dispersion (at the 5% level) in both green and brown industries. This result highlights the harmonizing

role of institutional investors in reducing disagreement across ESG rating providers, irrespective of sectoral environmental intensity. Family ownership is positively associated with the external rating dispersion (significant at the 10% level) in brown industries only. This finding suggests greater heterogeneity in disclosure strategies among family-controlled firms operating in environmentally intensive sectors, where incentives to manage or selectively frame environmental information may be powerful. Other governance variables, including board independence, board size, CEO duality, and board gender diversity, do not exhibit statistically significant effects in these specifications, indicating that sectoral differences dominate governance mechanisms in shaping cross-rater disagreement.

■ *Political environment*

The political context matters primarily for brown industries. Firms located in Democratic-governed states (Allblue) exhibit significantly higher dispersion (significant at the 5% level) in high-emission sectors. This result may reflect several mechanisms. Stricter environmental regulation and enforcement in Democratic-governed states may increase disclosure requirements and litigation risk, encouraging firms to adopt heterogeneous compliance and communication strategies. At the same time, rating agencies may differ in how they evaluate transition risk, regulatory exposure, and forward-looking environmental commitments. These factors can amplify interpretive divergence across providers in emissions-intensive sectors. No statistically significant effect of political affiliation is observed in green industries.

■ *Discussion and implications*

Overall, the comparison between green and brown industries reveals both common and sector-specific drivers of environmental rating dispersion. Persistence and the disciplining role of institutional ownership are robust across sectors, whereas firm size, growth opportunities, leverage, family ownership, and the political environment matter in more targeted ways. Regarding the EI of the explanatory variables, profitability (-13%) and institutional ownership (-10%, -14%) exhibit strong effects, suggesting that these are critical factors explaining the external discrepancy. Interestingly, the extent of family ownership has different effects on EI in green (-5%) and brown industries (+19%), with the latter showing the stronger economic impact.

From an energy policy perspective, these findings underscore that improving the comparability of environmental ratings requires sector-sensitive approaches. In particular, emissions-intensive industries appear more exposed to political and ownership-related sources of disagreement, highlighting the importance of more explicit disclosure guidance and greater alignment in rating methodologies for brown sectors.

4.2. Internal Dispersion

This section reports regression results for the U.S. 2002-2019 sample, distinguishing between green and brown industries based on unscaled greenhouse gas emissions. The dependent variable is internal environmental rating dispersion, defined as disagreement across environmental dimensions within a single rating provider's assessment. Positive (negative) coefficients indicate higher (lower) dispersion. Statistical significance is denoted by *, **, and *** for the 10%, 5%, and 1% levels, respectively. The regression results are in Table 4. We also report the Economic Impact (EI) measure of each independent variable in Table 4 and Figure 2.

Table 4

Associations between the internal dispersion of E measures

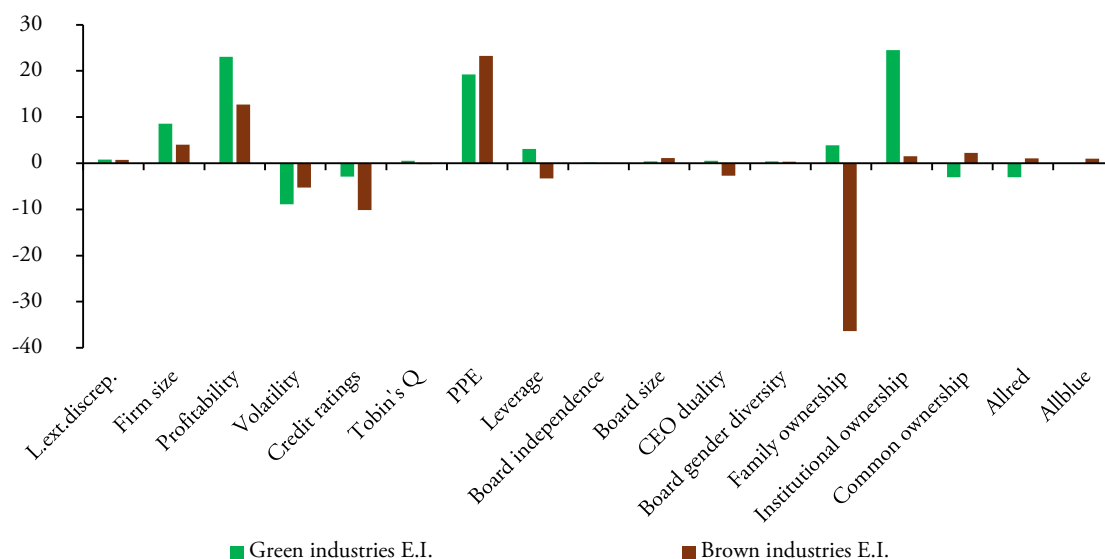
Elements	Green industries	EI (%)	Brown industries	EI (%)
Lagged internal dispersion	+ ***	0.78	+ ***	0.73
Firm size	+ ***	8.60	+ ***	4.03
Profitability	+ **	23.08	+	12.7
Volatility	-	-8.88	-	-5.24
Credit ratings	-	-2.89	-***	-10.15
Tobin's Q	+	0.55	-	-0.17
PPE	+ *	19.21	+***	23.28
Leverage	+	3.08	-	-3.32
Board independence	+	0.12	+	0.01
Board size	+	0.37	+*	1.14
CEO duality	+	0.55	-	-2.7
Board gender diversity	+ ***	0.39	+**	0.31
Family ownership	+	3.91	-**	-36.35
Institutional ownership	+ **	24.49	+	1.53
Common ownership	-	-3.04	+	2.22
Allred	-	-3.00	+	1.08
Allblue	-	-0.14	+	0.99
No. of observations	5,112	5,112	3,931	3,931

Note: This Table presents associations between the internal dispersion of E measures (*i.e.*, dispersion within a single data provider), measured by the logarithm of the standard deviation of those measures, and firm characteristics. Robust standard errors are estimated. The signs of the coefficients are reported. Standard errors are clustered at the firm level. ***, ** and * denote statistical significance at 1%, 5% and 10% level. EI is the Economic Impact measure.

■ Persistence of internal dispersion

Lagged internal discrepancy is positive and highly significant (at the 1% level) in both green and brown industries. This finding indicates strong persistence in within-rater disagreement

Figure 2

Internal Dispersion: Economic Impact

over time, suggesting that internal dispersion reflects structural features of firms and rating methodologies rather than short-run fluctuations.

■ *Firm characteristics*

Firm size is positively and strongly associated with internal dispersion (statistically significant at the 1% level) in both green and brown industries. Larger firms tend to operate across more heterogeneous activities and production processes, which naturally generates greater variation across environmental dimensions within a single rater's assessment. Profitability is positively associated with internal dispersion (at the 5% level) in green industries, but not in brown industries. This pattern suggests that, in environmentally less intensive sectors, more profitable firms exhibit greater differentiation across environmental pillars, possibly reflecting selective strengths and weaknesses across distinct dimensions of environmental performance.

Credit ratings are negatively and strongly associated with internal dispersion (significant at the 1% level) in brown industries only. For emissions-intensive sectors, higher credit quality appears to be associated with more internally consistent environmental assessments, potentially reflecting standardized processes or clearer environmental risk management.

Property, plant, and equipment (PPE) is positively associated with internal dispersion in both green (at the 10% level) and brown (at the 1% level) industries. Asset-intensive firms are likely

to face uneven environmental profiles across dimensions such as emissions, resource use, and operational efficiency, increasing within-rater dispersion.

■ *Governance and ownership structure*

Board size is positively associated with internal dispersion (significant at the 10% level) in brown industries, suggesting that more complex governance structures may correspond to greater heterogeneity in environmental practices across dimensions in emissions-intensive sectors.

Board gender diversity is positively and strongly associated with internal dispersion in both green (significant at the 1% level) and brown (significant at the 5% level) industries. This result is consistent with the idea that more diverse boards emphasize a broader set of environmental objectives, leading to stronger differentiation across environmental dimensions within a single rater's framework.

Family ownership is negatively associated with internal dispersion (significant at the 5% level) in brown industries. This finding suggests that family-controlled firms in emissions-intensive sectors tend to exhibit more internally uniform environmental profiles, possibly reflecting centralized control or more coherent environmental strategies.

Institutional ownership is positively associated with internal dispersion (significant at the 5% level) in green industries, whereas no significant effect is observed in brown industries. Institutional investors may encourage more granular, dimension-specific environmental reporting in cleaner sectors, thereby increasing within-rater differentiation across environmental pillars.

Regarding the EI of the explanatory variables, profitability (23%, 13%) is a relevant variable in explaining the extent of internal discrepancy, but its sign is opposite to that of external discrepancy. Tangibility (PPE) has a meaningful impact (19%, 23%), whereas institutional ownership (24%) in green industries and family ownership (-36%) in brown industries are critical values explaining the internal discrepancy.

Overall, the results indicate that internal environmental rating dispersion is systematic and persistent, with firm size and asset tangibility emerging as key drivers across sectors. Governance and ownership characteristics play a more nuanced role: board diversity and institutional ownership are associated with greater internal differentiation, particularly in green industries, while credit quality and family ownership reduce dispersion in brown industries.

4.3. Comparison of External *versus* Internal Dispersion

Taken together, the results reveal clear and systematic differences in the dispersion of external and internal environmental ratings. First, while both measures exhibit strong persistence over time, their determinants differ. External dispersion is more closely related to factors that complicate cross-provider comparability, such as firm size, growth opportunities, ownership structure, and the political environment. These results suggest that differences in methodological choices, aggregation schemes, and interpretations of forward-looking environmental information primarily drive cross-rater disagreement.

By contrast, internal dispersion is more strongly associated with firm-level heterogeneity in underlying activities and governance. Tangibility, defined as property, plant, and equipment divided by total assets (PPE), board gender diversity, and, in green industries, institutional ownership, is consistently linked to greater internal dispersion, indicating that within-rater variation reflects genuine multidimensional differences in environmental performance rather than pure measurement noise. At the same time, factors such as credit quality and family ownership reduce internal dispersion in brown industries, pointing to more uniform environmental profiles in firms with centralized control or standardized risk management.

Second, ownership structure plays opposite roles across the two concepts. Institutional ownership tends to reduce external dispersion while increasing internal dispersion. This pattern suggests that institutional investors are associated with clearer, more standardized disclosure that improves cross-rater comparability, while simultaneously encouraging more granular, dimension-specific reporting that sharpens within-rater differentiation across environmental pillars.

Overall, the contrast between internal and external dispersion underscores that environmental rating disagreement is not a monolithic phenomenon. External dispersion reflects limits to comparability across rating providers, whereas internal dispersion captures informative multidimensionality in firms' environmental profiles. From a policy perspective, this distinction implies that efforts to improve ESG rating consistency should prioritize methodological harmonization across providers while preserving valuable within-rater information about trade-offs across environmental dimensions.

The reported associations should be interpreted with caution, as ownership structure, governance, and profitability may be endogenous to firms' disclosure quality and ESG reporting infrastructure. For example, institutional investors may self-select into firms with more standardized environmental disclosure, while higher profitability may proxy for investments in internal reporting systems rather than directly affecting rating dispersion. Accordingly, the results describe systematic equilibrium patterns in environmental rating disagreement rather than causal effects.

5. CONCLUSIONS

This paper asks a simple but important question: when environmental (E) ratings disagree, what does that disagreement actually mean? We distinguish between two types. External dispersion occurs when different rating agencies give the same company very different environmental scores. Internal dispersion occurs when a single rating agency shows mixed performance across different environmental areas, such as emissions, resource use, or innovation. Using U.S. data from 2002–2019 and comparing industries with relatively high and low emissions, the evidence shows that external dispersion is greater in brown, high-emission industries, mainly due to ownership patterns and political factors, which create comparability problems. In greener industries, differences across raters are more closely linked to disclosure quality and institutional oversight. Within a single rating, green sectors show more variation across environmental areas, while brown industries tend to have more uniform internal scores.

We reach three main conclusions. First, disagreement in environmental ratings is not random or temporary. It is persistent over time, which means it reflects structural features of firms and rating systems. However, the source of disagreement matters. When ratings differ across agencies, the differences are often linked to company complexity, ownership structure, and political context. This result suggests that much cross-agency disagreement stems from differences in methodology, how agencies measure, weight, and interpret environmental information, rather than from fundamentally different views of the company's actual environmental performance.

Second, disagreement within a single rating often tells a different story. When a company scores well in some environmental areas but poorly in others, this frequently reflects real trade-offs in how the firm operates. Asset-heavy firms, firms with more diverse boards, and in some cases, firms with strong institutional investor presence tend to show greater variation across environmental dimensions. This kind of dispersion may contain useful information about the complexity of environmental performance rather than simple measurement error.

Third, ownership structure plays a key role, but in opposite ways depending on the type of dispersion. Institutional ownership tends to reduce disagreement across rating agencies, likely because it encourages clearer and more standardized disclosure. At the same time, it is associated with greater differentiation across environmental dimensions within a rating, suggesting more detailed and granular reporting. This contrast highlights why it is essential to distinguish between different forms of rating disagreement.

The policy implications are clear. Not all rating disagreements are a problem. Efforts to reduce dispersion risk mixing up two very different issues. Policymakers should focus on improving transparency and consistency across rating providers, and especially in emissions-intensive sectors, so that investors are not confused by avoidable methodological differences. At the same

time, regulation should avoid eliminating meaningful variation that reflects real environmental trade-offs within firms.

For investors, the message is equally important: a single ESG score does not tell the whole story. Understanding why ratings differ is crucial for sound capital allocation. For managers, clearer and more structured disclosure can reduce cross-agency disagreement while still reflecting environmental strengths and weaknesses.

Like all research, this study has limitations. Our classification of industries is based on emissions and focuses on U.S. firms. Future research could explore alternative definitions of environmental intensity and other regulatory contexts. It would also be valuable to examine how markets respond differently to external versus internal rating disagreement.

In short, disagreement in environmental ratings is not automatically a flaw. Sometimes it reflects inconsistent measurement; sometimes it reflects real complexity. Recognizing the difference is essential for building more credible ESG frameworks, designing better regulation, and making more informed investment and corporate decisions.

REFERENCES

AVRAMOV, D., CHENG, S., LIOUI, A., TARELLI, A. (2022). Sustainable investing with ESG rating uncertainty. *Journal of Financial Economics*, 145, 2, Part B, 642-664.

BERG, F., KOELBEL, J., and RIGOBON, R. (2022). Aggregate confusion: The divergence of ESG ratings. *Review of Finance* 26(6), 1315–1344.

BILLIO, M., COSTOLA, M., HRISTOVA, I., LATINO, C., & PELIZZON, L. (2021). Inside the ESG ratings: (Dis)agreement and performance. *Corporate Social Responsibility and Environmental Management*, 28(5), 1426–1445.

CHATTERJI, A., DURAND, R., LEVINE, D., and TOUBOUL, S. (2016). Do ratings of firms converge?. Implications for managers, investors, and strategy researchers. *Strategic Management Journal*, 37(8), 1597-1614.

CHRISTENSEN, D. M., SERAFEIM, G., and SIKOCHI, A. (2022). Why is Corporate Virtue in the Eye of the Beholder? The Case of ESG Ratings. *The Accounting Review*, 97(1), 147–175.

DI GIULI, A., and KOSTOVETSKY, L., (2014). Are red or blue companies more likely to go green? Politics and corporate social responsibility. *Journal of Financial Economics*, 111, 158–180.

- EDMANS, A. (2023). The end of ESG. *Financial Management*, 52, 3–17.
- GIBSON, R., GLOSSNER, S., KRUEGER, P., MATOS, P., STEFFEN, T. (2022). Do Responsible Investors Invest Responsibly? *Review of Finance*, 26(6), 1389–1432.
- GIBSON, R., KRUEGER, P., and SCHMIDT, P. S. (2021). ESG rating disagreement and stock returns. *Financial Analysts Journal*, 77, 104–127.
- HUBER, B. M., and COMSTOCK, M. (2017). *ESG Reports and Ratings: What They Are, Why They Matter*. Harvard Law School Forum on Corporate Governance and Financial Regulation.
- KOTSANTONIS, S., and SERAFEIM, G. (2019). Four Things No One Will Tell You About ESG Data. *Journal of Applied Corporate Finance*, 31(2), 50–59.
- PEDERSEN, L. H., FITZGIBBONS, S., and POMORSKI, L. (2021). Responsible investing: The ESG-efficient frontier. *Journal of Financial Economics*, 142(2), 572–597.
- PEÑA, J. I., DESENDER, K. A., and GUO, B. (2025). *Divergence and Uncertainty in ESG Markets: Metrics, Economic Policy, and Financial Implications*. Estudios de la Fundación. Serie Economía y Sociedad. Funcas.
- RAGHUNATHAN, T. (2003). An Approximate Test for Homogeneity of Correlated Correlation Coefficients. *Quality & Quantity*, 37, 99–110.
- SAUTNER, Z., VAN LENT, L., VILKOV, G., and ZHANG, R. (2023a). Firm-level climate change exposure. *The Journal of Finance*, 78(3), 1449–1498.
- SAUTNER, Z., VAN LENT, L., VILKOV, G., and ZHANG, R. (2023b). *Data for 'Firm-level Climate Change Exposure*. <https://doi.org/10.17605/OSF.IO/FD6JQ>
- SERAPEIM, G., YOON, A. (2023). Stock price reactions to ESG news: the role of ESG ratings and disagreement. *Review of Accounting Studies*, 28, 1500–1530.
- WONG, C., NELSON, J., and BRACKLEY, A. (2023). *Rate the Raters 2023: Expert Views on ESG Ratings*. <https://www.sustainability.com/globalassets/sustainability.com/thinking/pdfs/2023/rate-the-raters-report-april-2023.pdf>
- ZOELLER, T. (2025). Homogeneity test for multiple correlation coefficients (<https://www.mathworks.com/matlabcentral/fileexchange/25367-homogeneity-test-for-multiple-correlation-coefficients>), MATLAB Central File Exchange.

Invertir en la transición energética bajo incertidumbre profunda: modelos de planificación y métodos robustos

Antonio F. Rodríguez-Matas*

Resumen

Las inversiones necesarias para la transición energética son intensivas en capital, de larga vida útil y están expuestas a incertidumbres: evolución tecnológica, precios de combustibles, evolución de la demanda, aceptación social, regulación y *shocks* geopolíticos o climáticos. En este contexto, la planificación no puede basarse solo en escenarios centrales ni en hipótesis sobre un futuro conocido. Este artículo presenta un conjunto de herramientas desarrolladas para mejorar la planificación estratégica bajo incertidumbre profunda. En primer lugar, se describe el papel de los modelos de optimización de sistemas energéticos a largo plazo y se introduce openMASTER, un modelo abierto implementado en Python-Pyomo que representa con detalle tecnologías, sectores de demanda y restricciones técnicas y de política pública. En segundo lugar, se exponen tres desarrollos metodológicos que complementan el modelado: (i) un enfoque híbrido que combina optimización robusta y criterios de *minimax* del arrepentimiento para tratar de forma diferenciada las preferencias de decisión frente a incertidumbres; (ii) un método para integrar las correlaciones entre parámetros inciertos mediante conjuntos de incertidumbre basados en análisis de componentes principales, evitando recomendaciones sesgadas; y (iii) un marco para diseñar paquetes de políticas robustos sobre múltiples métricas, basado en exploración de escenarios, un algoritmo de selección iterativa y técnicas de interpretabilidad para identificar palancas y vulnerabilidades. Este artículo trata de analizar cómo estas herramientas permiten estresar carteras de inversión, identificar inversiones estructuralmente necesarias y diseñar políticas monitorizables y adaptables que reduzcan la probabilidad de arrepentimiento en inversiones de largo plazo.

Palabras clave: planificación energética, inversión, incertidumbre profunda, robustez, optimización, paquetes de políticas, modelos energéticos.

1. INTRODUCCIÓN

La transición energética se juega, en gran medida, en el terreno de las inversiones. A diferencia de otras políticas públicas, donde los efectos pueden ser relativamente reversibles, muchas decisiones energéticas tienen un fuerte componente de irreversibilidad: la vida útil de una red, una planta de generación, una infraestructura de almacenamiento o un parque de edificios rehabilitados se mide en décadas. Además, estas decisiones se toman en un contexto en el que los marcos regulatorios y los objetivos climáticos evolucionan, y donde tecnologías y mercados cambian a ritmos que son difíciles de anticipar.

Esta combinación —capital intensivo, horizontes largos e incertidumbre— plantea un reto muy importante: ¿cómo decidir hoy las inversiones que sean compatibles con los objetivos

* Instituto de Investigación Tecnológica (IIT), Universidad Pontificia Comillas.

de descarbonización y, a la vez, robustas frente a futuros adversos? La respuesta habitual ha sido analizar unos pocos escenarios contruidos en base a una combinación de hipótesis sobre las incertidumbres: “alto crecimiento económico”, “bajo coste de renovables”, “alto precio de carbono”, etc. Sin embargo, el sector energético actual enfrenta un tipo de incertidumbre más exigente. No se trata solo de dispersión alrededor de una expectativa, sino de “incertidumbres epistémicas”: situaciones en las que no hay consenso razonable sobre modelos causales, distribuciones de probabilidad o incluso sobre qué futuros deben considerarse plausibles. En este contexto, el valor de una herramienta no reside en predecir un futuro, sino en reducir la fragilidad de las decisiones ante sorpresas.

Los modelos de planificación energética a largo plazo ofrecen un marco cuantitativo para traducir objetivos generales (neutralidad climática, seguridad de suministro, asequibilidad) en necesidades coherentes de infraestructura y tecnología. No obstante, si se usan de forma determinista, estos modelos pueden dar una falsa sensación de precisión: una única trayectoria óptima puede ocultar vulnerabilidades relevantes. Este artículo parte precisamente de ese diagnóstico: mejorar cómo los modelos energéticos pueden apoyar decisiones cuando el futuro no es “probabilizable” ni reducible a unos pocos escenarios. Para ello, se sintetizan cuatro contribuciones científicas recientes: el desarrollo del modelo openMASTER (Rodríguez-Matas *et al.*, 2024), un método para mejorar la robustez ante incertidumbre epistémica (Rodríguez-Matas *et al.*, 2024), la integración de correlaciones entre incertidumbres (Rodríguez-Matas *et al.*, 2025) y un marco para diseñar paquetes de políticas robustos bajo incertidumbre profunda (Rodríguez-Matas *et al.*, 2026).

El texto se organiza en torno a un objetivo: del modelo a la decisión. Primero, se explica qué aportan los modelos de optimización de sistemas energéticos a las decisiones de inversión y se presenta openMASTER, un modelo “open source”¹ desarrollado con ese propósito. Segundo, se introduce el problema de la incertidumbre profunda en la transición y por qué obliga a complementar el análisis “óptimo” con criterios de robustez. Tercero, se describen tres desarrollos metodológicos que, sobre la base de openMASTER, permiten someter carteras y políticas a pruebas de estrés, incorporar relaciones realistas entre incertidumbres y diseñar paquetes de políticas robustos sobre múltiples objetivos. Finalmente, se discute cómo integrar estos resultados en procesos reales de planificación e inversión, tanto públicos como privados.

1 Un modelo *open-source* es aquel cuyo código fuente se publica con una licencia abierta que permite su uso, inspección, modificación y redistribución (sujeto a las condiciones de la licencia). En el contexto de la planificación energética, esto facilita la reproducibilidad (replicar resultados a partir de los mismos datos y supuestos), la auditoría de hipótesis y formulaciones, y la extensibilidad por terceros (por ejemplo, añadir sectores, restricciones o métricas), reduciendo además la dependencia de herramientas propietarias. En cualquier caso, la apertura del código no garantiza por sí misma la calidad del modelo: la credibilidad depende también de la documentación, los datos empleados, la verificación/validación y la trazabilidad de los escenarios analizados.

El objetivo no es presentar un recetario cerrado ni un resultado único, sino mostrar una forma de trabajo replicable: formular decisiones de inversión relevantes, representar el sistema con un modelo transparente y explorar sistemáticamente cómo diferentes futuros condicionan el desempeño de estrategias alternativas. En términos prácticos, se trata de reducir la probabilidad de arrepentimiento (*regret*) y de evitar inversiones frágiles que funcionen bien en un futuro central, pero fallen en otros escenarios adversos.

En la práctica, esta problemática se manifiesta tanto en decisiones públicas como privadas. Un operador de red debe decidir hoy refuerzos que habiliten la electrificación y el despliegue de renovables, pero su rentabilidad dependerá de patrones futuros de demanda, de la localización de renovables, y de la evolución de almacenamiento y flexibilidad. Una empresa industrial puede decidir inversiones en electrificación, captura de CO₂ o hidrógeno, pero su competitividad dependerá de los precios de electricidad, la disponibilidad del recurso energético, la regulación de emisiones y las cadenas de suministro. Incluso decisiones aparentemente menores, como ritmos de rehabilitación de edificios o despliegue de puntos de recarga, pueden generar *lock in*² si se adoptan estándares o infraestructuras que resulten incompatibles con una transición posterior más ambiciosa.

En este tipo de decisiones, el valor de una herramienta de planificación no se mide por su precisión predictiva, sino por su capacidad para reducir incertidumbre en el sentido operativo: (i) acotar el rango de necesidades estructurales, (ii) hacer explícitos los supuestos que sustentan una recomendación, y (iii) identificar condiciones bajo las cuales una estrategia dejaría de ser aceptable. Esta última idea es especialmente importante: en incertidumbre profunda, la planificación debe acompañarse de un sistema de seguimiento que permita corregir el rumbo. Este trabajo vuelve a esta idea en la sección 6, conectando resultados de modelización con mecanismos de monitorización y adaptación (Haasnoot *et al.*, 2013; Marchau *et al.*, 2019).

Como regla general, cuanto mayor sea la irreversibilidad y el horizonte de la decisión, menor debe ser el peso de la optimización puntual y mayor el de la robustez. Esto no significa renunciar al análisis cuantitativo: significa usarlo para construir margen de maniobra. En la transición energética, ese margen se materializa en inversiones que habilitan opciones (flexibilidad, redes, capacidades modulares), en políticas que pueden ajustarse sin colapsar el sistema y en mecanismos de seguimiento que detecten temprano desviaciones. La combinación de openMASTER y los métodos presentados busca precisamente dotar a los decisores de ese mapa de opciones, con rigor cuantitativo y con transparencia suficiente para que el resultado sea discutible y, por tanto, mejorable.

2 *Lock-in* (dependencia de trayectoria) es el fenómeno por el que decisiones de inversión irreversibles (costes hundidos y vida útil larga) generan inercia tecnológica e institucional, reduciendo el espacio de opciones futuras y elevando el coste de cambiar de senda, incluso si nueva información revela que la trayectoria inicial es subóptima o incompatible con objetivos posteriores; en el extremo, puede traducirse en activos varados (*stranded assets*).

2. MODELOS DE PLANIFICACIÓN ENERGÉTICA Y DECISIONES DE INVERSIÓN

Los modelos de planificación energética a largo plazo —en particular, los modelos de optimización de sistemas energéticos— representan el sistema como una red de tecnologías y flujos energéticos capaces de satisfacer demandas de servicios energéticos (movilidad, confort térmico, producción industrial, etc.) bajo restricciones técnicas y de política pública. Su utilidad, desde la perspectiva de la inversión, radica en que permiten analizar decisiones de forma sistémica: no como proyectos aislados, sino como piezas de un sistema con fuertes interdependencias.

En términos simples, un modelo de optimización estratégica busca la combinación de inversiones y operación que minimiza un criterio (por ejemplo, el coste total del sistema) sujeto a cumplir demandas y restricciones (por ejemplo, un presupuesto de carbono). Esta estructura facilita responder preguntas típicas de inversión: ¿qué tecnologías se despliegan cuando se endurece un objetivo de emisiones?, ¿qué cuellos de botella aparecen si se limita la expansión de redes o almacenamiento?, ¿qué combinación de electrificación, eficiencia y combustibles renovables reduce costes totales?, ¿cuándo conviene invertir en infraestructuras habilitadoras que no reducen emisiones por sí solas, pero permiten integrar renovables o electrificar demanda final?

Para un decisor, el valor añadido del modelado no se limita a la cifra final de coste o emisiones. Hay tres tipos de resultados especialmente relevantes para la inversión:

- *Necesidades de capacidad e infraestructura.* Muchos debates se formulan en términos de objetivos (porcentaje renovable, penetración de vehículos eléctricos, reducción de demanda). Un modelo traduce esos objetivos en cantidades de inversión coherentes (GW de renovables, capacidad de almacenamiento, refuerzos de red, equipos en edificios, etc.) y, crucialmente, en su secuenciación temporal. Esto es fundamental porque en transición energética no solo importa “dónde llegamos”, sino “cómo llegamos”: la infraestructura que se instala en 2030 condiciona el espacio de opciones en 2040.
- *Interacciones entre sectores.* La electrificación del transporte aumenta la demanda eléctrica y puede exigir inversión adicional en generación y redes; la descarbonización industrial puede demandar hidrógeno, que a su vez requiere electricidad y agua; y la eficiencia en edificios puede reducir picos de demanda y aliviar necesidades de capacidad. Un análisis sectorial aislado tiende a ignorar estas retroalimentaciones.
- *Comparación de carteras.* Una cartera de inversión rara vez se basa en una única tecnología. Lo habitual es decidir entre combinaciones: más solar + almacenamiento vs. más eólica + redes; electrificación directa vs. combustibles renovables; mayor rehabilitación vs. mayor expansión

de renovables. Un modelo permite comparar estas carteras bajo supuestos coherentes y cuantificar *trade offs*.

Ahora bien, los modelos también tienen limitaciones importantes que deben entenderse para no sobreinterpretar los resultados. En modelos de optimización estratégica, la representación es típicamente agregada (por ejemplo, por sectores o regiones), la competencia imperfecta y los fallos de mercado no siempre se modelan explícitamente, y es frecuente asumir coordinación y previsión perfectas (*perfect foresight*) en el horizonte temporal³. Estos supuestos son adecuados para estudiar necesidades estructurales y coherencias sistémicas, pero pueden distanciarse de cómo invierten realmente empresas y hogares bajo incertidumbre, restricciones financieras, fricciones regulatorias o preferencias heterogéneas.

Por ello, una práctica recomendada en planificación es usar el modelo como “laboratorio”: no para pronosticar, sino para explorar “qué tendría que ser cierto” para que una estrategia funcionase o fracasase. Esto requiere ampliar el análisis más allá de una única optimización determinista. El resto del artículo se centra en cómo hacerlo: primero con un modelo diseñado para ser transparente y extensible (openMASTER) y, después, con metodologías explícitas de robustez y exploración de escenarios.

2.1. Una breve tipología sobre modelos

Conviene distinguir, sin entrar en tecnicismos, tres familias de herramientas que suelen confluir en debates de inversión:

- Modelos de balance o contabilidad energética, que verifican coherencia física y pueden ser útiles para hojas de ruta, pero no son útiles para analizar inversiones óptimas.
- Modelos de optimización de sistemas, como los de tipo *bottom up*⁴, que determinan inversiones y operación al mínimo coste bajo restricciones. Son especialmente valiosos cuando se buscan carteras coherentes y cuando las interdependencias son fuertes (Pfenninger *et al.*, 2014).
- Modelos de simulación o de agentes, que pueden representar decisiones descentralizadas, fricciones y dinámicas de adopción, pero a costa de mayores requerimientos de parametrización y, a menudo, menor trazabilidad para entender por qué un resultado es óptimo o no.

³ Por coordinación perfecta se entiende que el modelo actúa como si las decisiones de inversión y operación de todos los agentes estuvieran alineadas y fueran consistentes (sin desajustes ni conflictos de incentivos). Por previsión perfecta (*perfect foresight* en inglés), que el decisor conoce *ex ante* la evolución futura de variables clave durante todo el horizonte de planificación (costes, demanda, precios, políticas), en lugar de decidir secuencialmente con información parcial.

⁴ En modelos *bottom-up*, el sistema se representa “desde abajo”: con tecnologías y procesos explícitos (rendimientos, costes, límites y sustituciones) y, a partir de esa desagregación, el modelo determina la combinación de inversiones y operación que cumple la demanda y las restricciones al mínimo coste.

Este artículo se sitúa en la segunda familia, pero incorpora elementos para acercarse a preocupaciones típicas de la tercera (por ejemplo, dinámicas de adopción y restricciones de comportamiento en la demanda) dentro de un marco de optimización, buscando un equilibrio entre detalle y manejabilidad computacional.

2.2. De la optimización a la utilidad: el papel de la exploración

Una crítica recurrente a los modelos de optimización es que, al minimizar coste, pueden producir transiciones demasiado rápidas o poco realistas. La literatura ha señalado que la optimización puede aproximar mal transiciones observadas si se interpreta literalmente como predicción (Trutnevyte, 2016). Sin embargo, esa misma literatura también resalta que estos modelos son muy útiles si se emplean como herramientas de exploración: sirven para identificar tensiones, cuellos de botella y condiciones de viabilidad.

Por tanto, el uso responsable no es elegir entre “optimización” o “realismo”, sino combinar la optimización con un análisis sistemático de incertidumbre y con un diálogo sobre supuestos. En la práctica, esto implica: (i) ejecutar el modelo bajo múltiples futuros plausibles, (ii) comparar carteras de inversión bajo métricas relevantes, y (iii) extraer reglas prácticas (por ejemplo, inversiones habilitadoras robustas) en lugar de fijarse únicamente en la trayectoria de un escenario.

3. OPENMASTER: UN MODELO *OPEN SOURCE* PARA PLANIFICACIÓN ESTRATÉGICA

openMASTER es un modelo de optimización energética de largo plazo desarrollado con el objetivo de analizar políticas y trayectorias tecnológicas de manera integrada y reproducible (Rodríguez-Matas *et al.*, 2024). Su diseño sigue una lógica *bottom up* y de equilibrio parcial: representa tecnologías y procesos de forma explícita, y determina inversiones y operación que satisfacen demandas exógenas de servicios energéticos al mínimo coste agregado, sujeto a restricciones técnicas y de política pública. Está implementado en Python-Pyomo y puede resolverse con *solvers* comerciales u *open source*, lo que facilita su adopción y extensión por parte de terceros⁵.

Desde la perspectiva de inversión y política, el interés de openMASTER radica en dos elementos: (i) una representación detallada del “lado de la demanda”, formulada en términos de servicios energéticos, y (ii) la modularidad para integrar dimensiones que suelen quedar fuera de modelos estratégicos convencionales.

5 Para descargar el modelo openMASTER: <https://github.com/IIT-EnergySystemModels/openMASTER>

3.1. Demanda como servicios, no solo como energía final

Formular la demanda en términos de servicios (kilómetros pasajero, confort térmico, toneladas de acero, etc.) permite representar explícitamente que la transición no es solo un cambio de oferta, sino también de tecnologías de uso final y de comportamiento. Por ejemplo, la misma demanda de movilidad puede satisfacerse con diferentes combinaciones de modos (vehículo privado, transporte público, movilidad activa) y tecnologías (ICE, BEV, etc.)⁶, con implicaciones distintas en inversión de infraestructura y en consumo energético. Del mismo modo, la demanda de calefacción puede satisfacerse con calderas, bombas de calor u otras tecnologías, y puede reducirse mediante mejoras de eficiencia, rehabilitación de edificios o cambios de hábitos.

Esta formulación es especialmente útil cuando se quieren analizar políticas que actúan sobre comportamientos o sobre el parque de equipos (por ejemplo, estándares de eficiencia, programas de rehabilitación, teletrabajo), ya que permite representar las consecuencias sistémicas de esas medidas.

3.2. Modularidad y extensiones relevantes para inversión

openMASTER se concibió como un marco extensible. Entre las funcionalidades incorporadas destacan:

- Representación endógena del comportamiento y elección en ciertas decisiones del lado de la demanda, permitiendo capturar que la adopción tecnológica no es automática ni instantánea, y que puede depender de costes generalizados, disponibilidad y restricciones.
- *Vintages*⁷ tecnológicos, que permiten representar la coexistencia de equipos antiguos y nuevos, los cambios de eficiencia, la dinámica de sustitución y los efectos de *lock in*.
- Circularidad y materiales, en la medida en que algunas decisiones de transición dependen de la disponibilidad de recursos, el reciclaje o las cadenas de suministro.
- Hibernación y reactivación de tecnologías, útil para representar capacidades que pueden quedar infrautilizadas temporalmente pero reactivarse si cambian condiciones.

⁶ ICE (*Internal Combustion Engine*) se refiere a vehículos con motor de combustión interna (gasolina o diésel). BEV (*Battery Electric Vehicle*) se refiere a vehículos 100 % eléctricos a batería, sin motor de combustión.

⁷ Por *vintages* tecnológicos se entiende una representación por cohortes de equipos según su año de instalación, de modo que el modelo distingue el parque “existente” de los “nuevos” y puede capturar vida útil, retiros, tasas de sustitución, y eficiencias/costes distintos por generación tecnológica, haciendo explícitos los efectos de inercia y *lock-in*.

- Restricciones de fiabilidad y adecuación, que permiten introducir métricas de seguridad de suministro más allá de los balances energéticos.

Estas dimensiones no son adornos técnicos, están directamente conectadas con decisiones de inversión. La transición energética se caracteriza por secuencias de inversión donde el tiempo importa: rehabilitar antes reduce necesidades de capacidad; invertir en redes habilita la electrificación; desplegar almacenamiento reduce los vertidos y facilita la integración renovable. Los modelos que solo capturan costes de generación sin representar demanda o restricciones de adecuación pueden subestimar inversiones habilitadoras o sobrestimar la facilidad de ciertas trayectorias.

3.3. Qué puede responder openMASTER (y qué no)

En la práctica, openMASTER puede utilizarse para:

- Identificar trayectorias tecnológicas coherentes con objetivos climáticos, incluyendo la secuenciación temporal de inversiones.
- Cuantificar *trade offs* entre costes, emisiones, dependencia energética y otros indicadores.
- Evaluar el efecto sistémico de políticas sectoriales y sus interacciones.
- Generar sendas monitorizables: series temporales de variables clave (capacidad, consumo, adopción tecnológica, etc.) que pueden usarse como referencia para seguimiento y ajuste.

Sin embargo, es crucial reconocer limitaciones para evitar un uso ingenuo:

- Previsión perfecta: como muchos modelos estratégicos, la versión base asume que el planificador conoce la trayectoria futura de las variables exógenas. Esto es útil para identificar necesidades estructurales, pero no representa la toma de decisiones secuencial con información parcial.
- Agregación y equilibrio parcial: el modelo no pretende reproducir la formación de precios en mercados complejos ni comportamientos estratégicos. Para análisis de mercado, se requieren modelos complementarios.
- Dependencia de datos y supuestos: la calidad de los resultados depende de supuestos sobre costes, potenciales y restricciones. La transparencia *open source* ayuda a auditar, pero no elimina del todo el problema.
- Complejidad computacional: incorporar más detalle sectorial y más dimensiones de incertidumbre incrementa el coste computacional; esto condiciona qué metodologías son factibles.

Estas limitaciones motivan la segunda parte del artículo: cómo complementar el modelado con metodologías que reconocen explícitamente la incertidumbre profunda y que permiten usar el modelo como laboratorio de estrés, no como “oráculo”.

3.4. Transparencia y reproducibilidad como activo para la decisión

La decisión de desarrollar openMASTER como modelo *open source* responde a un problema recurrente en planificación: la dificultad de auditar resultados cuando el modelo es propietario o no reproducible. En decisiones de inversión, especialmente públicas, esto genera dos riesgos: (i) dependencia de un proveedor único (*vendor lock in*) y (ii) pérdida de capacidad institucional para discutir supuestos y para adaptar el modelo a nuevas preguntas. Un modelo abierto permite que el debate se centre en hipótesis verificables (costes, potenciales, restricciones) y facilita la acumulación incremental de mejoras.

Además, la transparencia es crucial cuando se incorporan metodologías de robustez. Si una estrategia robusta tiene un coste mayor que el escenario central, es razonable que un decisor exija entender por qué: qué incertidumbres la justifican, qué riesgos reduce y cuánto. En modelos opacos, esa discusión es difícil. En modelos abiertos, es posible rastrear el origen de la “prima de robustez” en restricciones, parámetros o mecanismos concretos.

3.5. Resolución temporal y lectura de resultados para inversión

Un aspecto que a veces se pasa por alto es que el resultado más útil para apoyar la toma de decisiones en la inversión no es necesariamente el *mix* final, sino la dinámica: cuándo se despliegan capacidades, cuándo se sustituyen equipos y dónde aparecen picos de inversión. openMASTER opera en horizontes de largo plazo con discretización temporal (por ejemplo, periodos quinquenales) y puede incorporar representaciones simplificadas de operación dentro de cada periodo. Esta estructura es adecuada para planificación estratégica: permite capturar inversión acumulada y sustitución de *vintages* sin entrar en la complejidad de una representación horaria completo, que sería prohibitivo a escala multisectorial.

Para interpretar resultados de inversión, es útil distinguir tres capas:

- Capa de capacidad: inversión en GW de generación, km de redes (si se representa), capacidad de almacenamiento, y equipamiento en usos finales.
- Capa de flujos: consumo de vectores energéticos (electricidad, combustibles, hidrógeno), emisiones asociadas y cambios en el *mix*.

- Capa de restricciones activas: qué límites son “vinculantes” (presupuesto de carbono, potencial renovable, ritmos máximos de despliegue, adecuación). Esta capa suele explicar por qué una inversión aparece en una trayectoria y es clave para diseñar políticas: si una tecnología se despliega porque el modelo necesita cumplir un presupuesto de carbono bajo límites de electrificación, la interpretación política es distinta que si se despliega por pura competitividad económica.

En una lógica robusta, además, interesa entender cómo estas capas se distribuyen en el conjunto de escenarios: qué inversiones aparecen en casi todos los futuros y cuáles solo en subconjuntos específicos. Esta lectura es la que permite separar inversiones habilitadoras de inversiones contingentes.

4. INCERTIDUMBRE PROFUNDA Y POR QUÉ LOS ESCENARIOS “CENTRALES” NO BASTAN

En planificación energética se habla de incertidumbre de forma habitual, pero no todas las incertidumbres son iguales ni requieren las mismas herramientas. Una distinción útil (y coherente con la literatura de decisión bajo incertidumbre profunda) es entre incertidumbre aleatoria o “probabilizable” e incertidumbre epistémica o profunda.

En el primer caso, existe conocimiento o evidencia suficiente para describir una distribución de probabilidad razonable o, al menos, acordar un rango con mayor densidad alrededor de un valor esperado. Este tipo de incertidumbre se puede manejar con técnicas estándar: análisis de sensibilidad, Monte Carlo con distribuciones, o incluso optimización estocástica.

En la incertidumbre profunda, en cambio, no hay consenso sobre (i) el modelo causal relevante, (ii) la distribución de probabilidad de las variables clave o (iii) el conjunto de futuros plausibles. Esto ocurre, por ejemplo, cuando la dinámica depende de cambios tecnológicos disruptivos, de la aceptación social, de la regulación futura, o de *shocks* geopolíticos.

Más allá de la definición formal, hay señales prácticas para reconocer incertidumbre profunda en el sector energético:

- Cuando los rangos dependen de decisiones políticas futuras (por ejemplo, diseño de mercados, permisos, aceptación social).
- Cuando la innovación puede alterar la frontera tecnológica (aprendizaje, sustitución de materiales, nuevas cadenas de suministro).
- Cuando existen riesgos de *shocks* correlacionados (geopolítica, clima, materias primas) que pueden mover simultáneamente varios parámetros.

- Cuando la propia estructura del sistema puede cambiar (por ejemplo, electrificación masiva, aparición de nuevos vectores energéticos).

Para decisiones de inversión, la implicación es clara: optimizar para un futuro esperado puede generar decisiones frágiles. La robustez se vuelve una propiedad valiosa: estrategias que no necesariamente son óptimas en un escenario, pero que funcionan de forma aceptable en un rango amplio de futuros.

La literatura de “Decision Making under Deep Uncertainty” (DMDU)⁸ ha formalizado este enfoque. En lugar de buscar la estrategia que maximiza la esperanza matemática, se prioriza identificar estrategias “satisfactorias” o que minimicen el peor resultado, el arrepentimiento máximo o la probabilidad de cruzar umbrales indeseables. Además, estos enfoques suelen combinar dos ingredientes: (i) exploración sistemática de futuros mediante grandes conjuntos de escenarios, y (ii) métodos de *scenario discovery*⁹ para identificar qué combinaciones de incertidumbres hacen a una estrategia vulnerable.

Desde un punto de vista operativo, esto encaja bien con el uso de modelos de optimización. Un modelo como openMASTER produce una trayectoria coherente para un conjunto de supuestos. Si repetimos el ejercicio sobre cientos o miles de combinaciones plausibles de incertidumbres, obtenemos una nube de resultados: distribución de costes, emisiones, dependencia, etc. La pregunta deja de ser “¿cuál es la trayectoria óptima?” y pasa a ser “¿qué estrategias y políticas siguen funcionando cuando el futuro cambia?”, “¿qué incertidumbres hacen colapsar una estrategia?”, y “¿qué inversiones aparecen de forma persistente como necesarias?”.

Lo relevante para inversión es que este enfoque permite separar tres categorías de decisiones:

- 1) Inversiones estructuralmente necesarias, robustas en casi todos los futuros (por ejemplo, ciertas infraestructuras habilitadoras).
- 2) Inversiones contingentes, que son valiosas solo si se materializan ciertas condiciones (por ejemplo, hidrógeno a gran escala si la electrificación directa es limitada o si ciertos sectores crecen).
- 3) Inversiones frágiles o con alto riesgo a la baja, que funcionan bien en un futuro favorable pero generan pérdidas grandes en futuros adversos.

⁸ *Decision Making under Deep Uncertainty* (DMDU) agrupa enfoques de apoyo a la decisión para contextos en los que no hay consenso sobre modelos, probabilidades o incluso el conjunto de futuros plausibles; en lugar de optimizar para un “escenario esperado”, prioriza estrategias robustas que mantengan un desempeño aceptable en muchos futuros.

⁹ *Scenario discovery* es un conjunto de técnicas (a menudo basadas en simulación y métodos estadísticos/ML) que se basa en la ejecución masiva de escenarios e identifica qué combinaciones de factores explican los casos en los que una estrategia falla o cruza umbrales, delimitando regiones del espacio de futuros asociadas a vulnerabilidad o éxito.

Este artículo muestra herramientas específicas para implementar esta lógica en modelos de optimización: cómo tratar distintas fuentes de incertidumbre de manera diferenciada, cómo incorporar correlaciones realistas y cómo diseñar paquetes de políticas robustos sobre múltiples métricas.

4.1. Criterios de decisión en incertidumbre profunda

En la práctica, la robustez puede implementarse con criterios distintos, según el apetito de riesgo y el tipo de decisión. Tres criterios aparecen de forma recurrente:

- Minimizar el peor caso (*minimax*): prioriza evitar resultados muy malos, aun a costa de sacrificar rendimiento en escenarios favorables.
- Minimizar el arrepentimiento máximo (*minimax regret*): busca la estrategia que minimiza el mayor arrepentimiento entre escenarios, donde el arrepentimiento en cada escenario es la diferencia de desempeño (por ejemplo, coste o utilidad) entre esa estrategia y la estrategia óptima para ese escenario si se conociera *ex ante* (Savage, 1951).
- Suficiencia con umbrales: define niveles mínimos aceptables (por ejemplo, emisiones máximas, dependencia máxima, coste máximo) y busca estrategias que cumplan en la mayoría de futuros, permitiendo *trade offs*¹⁰ explícitos.

Estos criterios se pueden utilizar y combinar en función del objetivo. Por ejemplo, la optimización robusta con presupuesto de incertidumbre captura un *minimax* “moderado” (Bertsimas & Sim, 2004), mientras que la evaluación con arrepentimiento permite comparar estrategias y cuantificar el coste de no haber sido perfectos *ex post* (Rodríguez-Matas *et al.*, 2024).

5. TRES HERRAMIENTAS PARA PLANIFICAR INVERSIONES ROBUSTAS

Esta sección presenta tres desarrollos metodológicos que, integrados con un modelo como openMASTER, permiten pasar de “optimización bajo un escenario” a “apoyo a la decisión bajo incertidumbre profunda”. Aunque cada método responde a un reto distinto, comparten un principio: estructurar decisiones para minimizar exposición a resultados extremos y, a la vez, extraer información práctica para la inversión (qué inversiones se activan, cuáles son contingentes, qué variables deben monitorizarse).

10 Por *trade-off* se entiende un compromiso: mejorar un objetivo suele implicar empeorar otro (por ejemplo, menos emisiones a cambio de mayor coste).

5.1. Tratar de forma diferenciada incertidumbres epistémicas

Una primera limitación de muchos análisis es que tratan todas las incertidumbres igual. Sin embargo, desde el punto de vista de un decisor, no es lo mismo la variabilidad “normal” de una demanda que la incertidumbre epistémica sobre el coste futuro de una tecnología emergente. En un caso, puede ser aceptable optimizar en torno a un valor esperado; en otro, puede ser preferible protegerse frente a resultados adversos porque la información es pobre y el error puede ser grande.

La optimización robusta ofrece un marco matemático para este tipo de protección: en lugar de suponer valores fijos, se define un conjunto de incertidumbre y se busca una solución que sea factible y razonablemente buena para cualquier realización dentro de ese conjunto. El enfoque de Bertsimas y Sim es especialmente útil porque introduce un “presupuesto de incertidumbre” que controla el grado de conservadurismo: permite que solo una fracción de parámetros tome su peor valor a la vez, lo que refleja que no todos los *shocks* negativos ocurren simultáneamente.

No obstante, el criterio puramente de “peor caso” puede ser excesivamente conservador si se aplica indiscriminadamente. Por ello, nuestro trabajo (Rodríguez-Matas *et al.*, 2024) propone un enfoque híbrido: combinar optimización robusta (para protegerse frente a incertidumbres epistémicas que afectan a la viabilidad del sistema) con un criterio de *minimax* del arrepentimiento (para evaluar el coste de no haber elegido la estrategia óptima *ex post* frente a incertidumbres económicas). El algoritmo de arrepentimiento compara, para cada escenario, el desempeño de una estrategia con el mejor desempeño que habría sido posible si se hubiese conocido ese escenario de antemano. Minimizar el arrepentimiento máximo ayuda a seleccionar estrategias que no pierdan demasiado frente a la mejor decisión con información perfecta.

Operativamente, el método sigue una lógica de “prueba de estrés” (*stress-test*):

- Se identifican incertidumbres clave y se clasifican según su naturaleza y relevancia para la decisión (aquellas que afectan a la viabilidad del sistema energético o a las que afectan al rendimiento económico).
- Para las incertidumbres que afectan a la viabilidad (por ejemplo, disponibilidad de recursos), se construye un conjunto de incertidumbre y se obtiene una estrategia robusta frente a peores escenarios con un nivel de protección controlado por el presupuesto.
- Para evaluar esa estrategia, y comparar alternativas, se calcula la decisión que minimiza el arrepentimiento sobre el conjunto de incertidumbres que afectan al rendimiento económico.

Desde la óptica de inversión, el resultado no es solo una solución robusta, sino una caracterización de *trade offs*: cuánto coste adicional (prima de robustez) se paga por reducir el riesgo a la baja, y qué incertidumbres son responsable de las mayores vulnerabilidades. Esta información es crucial para decidir si una inversión se justifica como un seguro frente a futuros adversos.

5.2. Integrar correlaciones

Una segunda limitación frecuente en análisis bajo incertidumbre es asumir independencia entre incertidumbres. Sin embargo, muchas incertidumbres energéticas están correlacionadas: los precios de combustibles fósiles tienden a moverse conjuntamente; los costes de tecnologías pueden compartir cuellos de botella (materiales, cadenas de suministro); y variables socioeconómicas pueden afectar simultáneamente a demanda residencial, transporte e industria, entre otras relaciones existentes en el sector energético.

Ignorar estas correlaciones puede sesgar la toma de decisiones. Por ejemplo, un análisis que permite que el precio del gas sea alto mientras el del crudo es bajo puede generar escenarios internamente inconsistentes y, por tanto, estrategias que parecen robustas pero lo son únicamente frente a futuros irreales. Asumir independencia puede subestimar el riesgo de *shocks* conjuntos.

Nuestro trabajo (Rodríguez-Matas *et al.*, 2025) incorpora un enfoque propuesto en la literatura reciente de optimización robusta basada en datos: construir conjuntos de incertidumbre a partir de datos históricos, pero de forma computacionalmente asequible, mediante el análisis de componentes principales (PCA). La idea es capturar la estructura dominante de correlación en un espacio reducido y construir un conjunto de incertidumbre poliédrico o elipsoidal que refleje esas dependencias. En la práctica:

- Se genera un conjunto de escenarios (por ejemplo, combinaciones de precios y costes) basado en datos o supuestos expertos.
- Se aplica PCA para identificar las direcciones principales de variación conjunta.
- Se define un conjunto de incertidumbre en el espacio transformado que preserve las correlaciones existentes.
- Se resuelve la optimización robusta con ese conjunto, obteniendo soluciones que se protegen frente a *shocks* correlacionados plausibles.

Para inversión, el mensaje es directo: un buen *stress test* debe ser realista en la estructura de los *shocks*. Invertir para protegerse frente a combinaciones imposibles puede ser ineficiente; no protegerse frente a *shocks* conjuntos plausibles puede ser peligroso.

5.3. Diseñar paquetes de políticas robustos sobre múltiples métricas

Las inversiones energéticas dependen de políticas. Además, para la transición energética, rara vez basta un único instrumento: la evidencia y la práctica apuntan a la necesidad de paquetes de políticas (*policy packages*) donde los instrumentos se refuerzan o compensan. Sin embargo, evaluar paquetes es complejo: el espacio de combinaciones crece exponencialmente y los instrumentos pueden tener efectos no lineales o generar vulnerabilidades nuevas (por ejemplo, una política que reduce CO₂ pero empeora contaminación local o incrementa costes de forma inaceptable bajo ciertos futuros), además de interactuar de forma compleja con las incertidumbres presentes en el sector energético.

Nuestro trabajo (Rodríguez-Matas *et al.*, 2026) propone un marco de apoyo a la decisión para diseñar paquetes robustos bajo incertidumbre profunda, combinando: (i) exploración de escenarios, (ii) selección iterativa de políticas, y (iii) técnicas de interpretabilidad para identificar palancas y vulnerabilidades.

El procedimiento se estructura en cuatro pasos:

- 1) Diseño experimental (XLRM): se definen incertidumbres (X), palancas de política (L), relaciones (R) representadas por un modelo energético (en nuestro caso openMASTER), y métricas de desempeño (M) como costes acumulados, emisiones, dependencia energética o contaminantes.
- 2) Preselección de políticas: a partir de un gran conjunto de escenarios (por ejemplo, un muestreo factorial de políticas e incertidumbres), se entrena un modelo de *machine learning* (*Random Forest*¹¹) para predecir métricas a partir de X y L. Las importancias de las variables se estiman con valores SHAP normalizados¹², lo que permite identificar qué políticas son más influyentes y reducir el espacio de políticas a analizar con mayor detalle, descartando aquellas que no tengan suficiente impacto en nuestras métricas objetivo.
- 3) Diseño iterativo del paquete robusto: se define un indicador de protección en el “peor caso” para cada política y métrica: cuánto mejora (o empeora) el peor resultado al activar una política. Estas protecciones se agregan con pesos que reflejan prioridades del decisor (por ejemplo, equilibrar coste, emisiones y dependencia). El algoritmo añade iterativamente la política que más mejora la protección agregada, recalculando en cada iteración las protecciones condicionadas a las políticas ya activas. Esto permite capturar interacciones: una política puede ser poco útil sola, pero valiosa en combinación.

11 Un *Random Forest* es un método de aprendizaje automático basado en un conjunto (ensamble) de árboles de decisión; al promediar sus predicciones, suele mejorar la precisión y reducir el sobreajuste frente a un único árbol.

12 Los valores SHAP asignan a cada variable una contribución a la predicción del modelo (en términos de teoría de valores de Shapley); normalizarlos permite expresar esas contribuciones en una escala comparable (por ejemplo, como proporciones) para ordenar la importancia relativa de políticas e incertidumbres.

- 4) Mapeo de sendas: una vez identificado el paquete robusto, se analizan las trayectorias tecnológicas, energéticas y de emisiones asociadas bajo la incertidumbre residual. Esto traduce un paquete de políticas abstracto en necesidades concretas de inversión (capacidad, electrificación, vectores energéticos) y permite definir qué variables monitorizar para detectar desviaciones y activar ajustes, conectando con marcos adaptativos como DAPP¹³.

Lo importante para inversión es que el resultado es doble: (i) un paquete de políticas que reduce riesgos extremos en varias métricas para generar un clima seguro para la inversión y (ii) un conjunto de trayectorias tecnológicas plausibles bajo ese paquete, que pueden guiar decisiones de inversión pública (redes, transporte, rehabilitación) y señales para inversión privada (mercados de tecnologías, infraestructura).

5.4. Cómo se complementan estas herramientas

Es tentador pensar que un método robusto resuelve el problema. Sin embargo, cada pieza atiende un fallo distinto del análisis convencional:

- El algoritmo híbrido “optimización robusta + *minimax* del arrepentimiento” aborda la tensión entre protección frente a incertidumbre epistémica y exceso de conservadurismo económico. Permite graduar la protección y medir su coste.
- El tratamiento de correlaciones evita que el *stress test* sea artificial. Es una condición de realismo: sin correlaciones plausibles, introducir robustez puede ser engañoso.
- El diseño de paquetes de políticas reconoce que, para inversión, el decisor no elige solo tecnologías, sino marcos de política que inducen inversiones privadas y coordinan sectores. Un paquete robusto reduce el riesgo de fallos sistémicos y facilita gobernanza.

En conjunto, estas herramientas permiten pasar de “decidir hoy una trayectoria” a “decidir hoy un conjunto de acciones y una forma de adaptación”. En términos de inversión, esto se traduce en carteras con componentes habilitadores (inversiones que amplían opciones futuras) y componentes contingentes (inversiones que se activan si ciertos hitos se materializan).

5.5. La “prima de robustez”

En análisis de inversión, una objeción recurrente a la robustez es que encarece el plan. Esto es cierto en sentido contable: protegerse frente a resultados adversos suele implicar invertir antes o

13 *Dynamic Adaptive Policy Pathways* (DAPP) es un enfoque de planificación adaptativa que diseña sendas de actuación por etapas, definiendo puntos de decisión y señales de seguimiento para ajustar, reforzar o cambiar las políticas a medida que se revela la evolución del sistema, con el objetivo de mantener la estrategia viable bajo incertidumbre profunda.

en tecnologías más flexibles, lo que puede aumentar el coste en escenarios favorables. La cuestión relevante es si ese sobre coste es razonable como seguro frente a pérdidas potenciales en escenarios adversos.

Una forma intuitiva de entenderlo es la siguiente. Supongamos que una cartera A es la más barata si los costes de renovables caen rápido y la demanda crece moderadamente, pero falla si los costes de almacenamiento no bajan y la demanda eléctrica crece más de lo esperado. Una cartera B invierte más en redes y flexibilidad desde el principio; en el escenario favorable es algo más cara, pero en el escenario adverso evita cortes de suministro o inversiones tardías muy costosas. La diferencia de costes en el escenario central es la prima de robustez: la diferencia en el peor caso es el valor del seguro. El algoritmo híbrido desarrollado en nuestro trabajo permite cuantificar ambos y discutir explícitamente el apetito de riesgo.

Cuando se aplica a un modelo como openMASTER, la decisión robusta no cambia la tecnología por capricho, típicamente desplaza decisiones hacia opciones que funcionan en un rango más amplio: tecnologías flexibles, inversiones habilitadoras y medidas que reducen exposición a *shocks* (por ejemplo, eficiencia o electrificación cuando reduce dependencia de combustibles volátiles). El análisis de arrepentimiento añade un matiz esencial: ayuda a identificar qué parte de la robustez proviene de evitar decisiones muy malas, no solo de ser conservador. Esto es particularmente relevante en inversiones con riesgo de *lock in*: una mala decisión inicial puede condicionar el sistema durante décadas.

Si esto no fuese suficientemente complejo, planificar utilizando combinaciones “cruzadas” poco plausibles (gas muy alto y petróleo muy bajo, o baterías muy baratas mientras materiales críticos son escasos), las estrategias que resultan “robustas” pueden serlo frente a futuros irreales, y frágiles frente a futuros plausibles. Por ello, el enfoque basado en PCA puede introducir sobre costes pero captura estas dependencias: no obliga a modelar todas las correlaciones una a una, sino que identifica patrones dominantes de variación conjunta y construye sobre ellos el conjunto de incertidumbre.

Por último, en el caso del diseño de paquetes de políticas, el procedimiento puede leerse como una forma de priorización informada. En términos de inversión, esta lógica se traduce en que un paquete robusto suele combinar: (i) señales de largo plazo que orientan la inversión privada (por ejemplo, una senda creíble de precio de carbono o estándares), (ii) inversiones públicas habilitadoras (redes, transporte, rehabilitación) y (iii) medidas de acompañamiento que reducen riesgos distributivos y de aceptación. El modelo permite cuantificar cómo estas piezas afectan a capacidades y flujos y, por tanto, a necesidades de inversión.

En definitiva, el marco puede utilizarse como herramienta de *due diligence*¹⁴ de escenarios para empresas e inversores. En lugar de preguntar “¿cuál será el precio de la electricidad en

14 *Due diligence* se refiere a un proceso de evaluación sistemática previa para identificar riesgos, supuestos críticos y condiciones de viabilidad de una inversión; aplicado a escenarios, implica contrastar el desempeño de una estrategia bajo múltiples futuros plausibles y detectar dónde se vuelve vulnerable.

2040?”, se puede preguntar “¿en qué rangos de precios y costes mi inversión sigue siendo competitiva?” y “¿qué combinaciones de factores me hacen vulnerable?”. La respuesta no es una predicción, sino un mapa de exposición: identifica qué variables explican el riesgo y qué estrategias de cobertura (contratos, flexibilidad tecnológica, diversificación) pueden reducirlo. Esta aproximación es coherente con la idea de que, en transición energética, el riesgo principal no es la volatilidad diaria, sino el riesgo de trayectoria (*path risk*): quedar atrapado en una senda tecnológica o regulatoria que luego se demuestra inviable o excesivamente costosa. Este riesgo se reduce priorizando decisiones robustas, aceptando cuando sea necesario una prima de robustez (un sobrecoste moderado hoy) a cambio de menor vulnerabilidad mañana.

6. DE LA HERRAMIENTA A LA DECISIÓN: CÓMO USAR ESTOS RESULTADOS EN INVERSIÓN

Para que estas herramientas sean útiles, deben integrarse en procesos reales de decisión. Un riesgo habitual es tratar el modelado como un ejercicio “de informe”: se ejecuta el modelo, se publica una cifra y se archiva. El enfoque robusto sugiere, en cambio, un ciclo iterativo similar a la gestión de riesgos: formular decisiones, explorar incertidumbres, identificar vulnerabilidades, diseñar medidas de mitigación y monitorizar.

A continuación se propone una guía práctica —no normativa— de cómo incorporar modelos y robustez en decisiones de inversión:

— Definir la decisión y el horizonte

Toda inversión relevante tiene una decisión concreta: priorizar redes vs. almacenamiento, acelerar rehabilitación vs. esperar, invertir en infraestructura de hidrógeno o en electrificación directa. La primera tarea es explicitar el horizonte (2030, 2040, 2050) y los criterios de desempeño (coste total, emisiones acumuladas, dependencia en 2050, etc.). El marco XLRM es útil para estructurar esta definición.

— Generar carteras candidatas

En lugar de comparar “miles de variables”, es preferible construir un conjunto de carteras o estrategias candidatas (por ejemplo, carteras de inversión o combinaciones de políticas) y evaluar su desempeño. Aquí openMASTER funciona como generador coherente de trayectorias, asegurando que cada cartera respeta balances y restricciones.

— Stress test sobre un espacio amplio de futuros

El análisis robusto sustituye el enfoque “escenario central + sensibilidad” por un conjunto amplio. Esto puede construirse con diseños factoriales, muestreos de rangos,

o conjuntos basados en datos. La clave es que el conjunto sea suficientemente diverso y que incluya correlaciones plausibles cuando sea relevante.

— *Medir robustez y trade offs*

Las decisiones de inversión no se reducen a un número. La metodología permite evaluar:

- Prima de robustez (coste adicional medio o en percentiles por reducir riesgos).
- Arrepentimiento máximo (cuánto se pierde frente a la decisión perfecta *ex post*).
- Protección ante el peor caso en múltiples métricas (coste, emisiones, dependencia, calidad del aire).

Estos indicadores ayudan a decidir cuánto “seguro” merece pagarse y en qué dimensiones.

— *Identificar vulnerabilidades y definir hitos*

Una ventaja de la exploración de escenarios es que permite traducir resultados a condiciones observables: por ejemplo, “la estrategia falla si los costes de electrificación no caen y la demanda industrial crece”. Estas condiciones se convierten en variables a monitorizar (hitos): evolución de costes, demanda, precios, aceptación social. En inversión, esto facilita cláusulas de ajuste, secuenciación y opciones reales: invertir hoy en infraestructura habilitadora y mantener la flexibilidad para escalar o desescalar según se revele el futuro.

— *Conectar con gobernanza y adaptación*

El mapeo de sendas convierte políticas en trayectorias monitorizables: capacidad renovable, electrificación, consumo de gas, despliegue de hidrógeno. Esto permite integrar el modelado en mecanismos de gobernanza: revisiones periódicas de planes, indicadores de seguimiento, y reglas de contingencia (por ejemplo, si el despliegue de bombas de calor se desvía de la trayectoria, activar medidas adicionales). Este enfoque se alinea con la idea de políticas adaptativas: diseñar no solo un plan, sino un plan con gatillos de adaptación.

En síntesis, el valor para invertir no consiste en ayudarte a “decidir una vez”, sino a mejorar la calidad del proceso: identificar inversiones habilitadoras, evitar apuestas frágiles y diseñar políticas que reduzcan riesgos extremos manteniendo flexibilidad.

Estas herramientas pueden servir a distintos actores, con preguntas diferentes:

- Administraciones y planificadores: identificar inversiones públicas habilitadoras, diseñar planes nacionales y regionales, evaluar coherencia entre objetivos y presupuestos.
- Reguladores: testear cómo señales regulatorias (por ejemplo, precios de carbono o estándares) afectan trayectorias y dónde se generan riesgos de inequidad o de aceptación.
- Empresas e inversores: evaluar robustez de carteras tecnológicas (por ejemplo, electrificación vs. hidrógeno) frente a rangos plausibles de precios y costes, y diseñar estrategias de cobertura.
- Sociedad civil y academia: auditar supuestos y participar en la deliberación sobre *trade offs*.

La utilidad práctica aumenta cuando los resultados se presentan como “reglas de decisión” y no solo como números: por ejemplo, “si el coste de X no cae por debajo de Y en 2030, conviene reforzar la palanca Z” o “la cartera A es robusta salvo en futuros con alta demanda industrial y bajo despliegue renovable, por lo que estos dos indicadores deben monitorizarse”. Este tipo de conclusiones es precisamente el valor añadido del enfoque de robustez y de la exploración de escenarios.

7. IMPLICACIONES PARA LA AGENDA DE INVERSIÓN

Aunque los resultados cuantitativos dependen del caso de estudio y de los supuestos, el enfoque aporta implicaciones generales para agendas de inversión en la transición:

- *Priorizar inversiones habilitadoras y de flexibilidad.* En muchos sistemas, la expansión renovable exige inversiones en redes, almacenamiento y gestión de demanda. Estas inversiones suelen aparecer como necesarias en múltiples trayectorias y, por tanto, tienden a ser robustas.
- *Tratar el lado de la demanda como cartera de inversión, no como residuo.* La electrificación y la eficiencia en edificios, movilidad e industria no son solo “resultados”, sino inversiones con dinámicas de adopción y *lock in*. Representarlas en términos de servicios energéticos ayuda a identificar dónde conviene invertir primero para reducir necesidades estructurales posteriores.
- *Diseñar políticas como paquetes con evaluación multimétrica.* Instrumentos aislados pueden generar vulnerabilidades. El análisis multimétrica permite anticipar efectos colaterales (por ejemplo, PM2.5 en escenarios de sustitución hacia biomasa) y diseñar paquetes que equilibren objetivos.

- *Incorporar correlaciones en stress tests.* En un entorno geopolítico y de cadenas de suministro tensas, los *shocks* conjuntos son plausibles. Integrar correlaciones evita recomendaciones basadas en futuros incoherentes.
- *Usar resultados para monitorización y adaptación.* El modelado puede informar qué variables deben seguirse para evaluar si la planificación va por buen camino y cuándo activar medidas adicionales, reduciendo el riesgo de *lock in*.

Estas implicaciones son particularmente relevantes para instituciones que planifican inversiones públicas (redes, transporte, rehabilitación) y para reguladores que deben dar señales estables a inversión privada. Un enfoque robusto no elimina la incertidumbre, pero ayuda a gestionarla explícitamente.

7.1. Inversión como secuencia y como opción

Una idea útil para conectar con finanzas de la inversión es la noción de opciones reales: invertir en activos que mantienen abiertas alternativas futuras puede tener valor incluso si no son estrictamente óptimos en el escenario central. Por ejemplo, reforzar redes y digitalización puede habilitar electrificación y flexibilidad; invertir en capacidades de almacenamiento modular puede reducir riesgo de vertidos si la penetración renovable se acelera; y programas de rehabilitación escalables pueden adaptarse a ritmos distintos de transición.

Los métodos presentados ayudan a identificar este tipo de inversiones al mostrar que aparecen en muchos futuros. También ayudan a identificar inversiones que, aunque rentables en un futuro favorable, pueden generar altos arrepentimientos en futuros adversos y, por tanto, deberían estructurarse como inversiones contingentes o escalonadas.

7.2. Implicaciones para políticas de inversión pública

En el ámbito público, estas herramientas son relevantes para el diseño de instrumentos que movilizan inversión privada (subastas, contratos por diferencia, incentivos fiscales, estándares). Un paquete robusto no es necesariamente el más agresivo, sino el que reduce vulnerabilidades y mantiene coherencia intersectorial. Por ejemplo, acelerar electrificación sin inversión en redes puede generar cuellos de botella y costes sistémicos; impulsar biocombustibles sin considerar la calidad del aire puede introducir externalidades; o apostar por un único vector (por ejemplo, hidrógeno) sin evaluar contingencias puede aumentar el riesgo si los costes no caen.

La evaluación multimétrica y el análisis de trayectorias ayudan a anticipar estos efectos y a diseñar instrumentos coordinados.

7.3. Conexión con marcos europeos de inversión

En Europa, buena parte de la inversión en transición está condicionada por marcos comunitarios (objetivos climáticos, regulación de redes, taxonomía, fondos y ayudas). En este contexto, las herramientas transparentes de planificación robusta pueden desempeñar un papel adicional: facilitar consistencia entre planes nacionales y marcos europeos, justificar necesidades de inversión con lógica sistémica y, sobre todo, proporcionar trazabilidad cuando se evalúan proyectos bajo múltiples objetivos (clima, seguridad, coste, resiliencia, competitividad y equidad). Además, la exigencia creciente de resiliencia frente a *shocks* (energéticos, climáticos, de cadenas de suministro) refuerza la necesidad de incorporar correlaciones y de diseñar paquetes de políticas que sean coherentes con distintos riesgos. El enfoque descrito puede contribuir a una gobernanza más robusta: no solo seleccionar proyectos, sino construir carteras de inversión que mantengan opciones abiertas y reduzcan vulnerabilidades estructurales.

8. CONCLUSIONES

La transición energética exige tomar decisiones de inversión en un contexto de incertidumbre profunda. En estas condiciones, la planificación basada en un único escenario o en una optimización determinista puede conducir a estrategias frágiles. El enfoque presentado en este artículo combina dos ingredientes: un modelo *open source* de optimización de sistemas energéticos (openMASTER) y tres metodologías complementarias de decisión robusta.

openMASTER aporta una representación tecnológica y sectorial detallada, formulada en términos de servicios energéticos y diseñada para ser transparente y extensible. Sobre esa base, las metodologías desarrolladas permiten: (i) tratar de forma diferenciada fuentes de incertidumbre y preferencias de decisión mediante optimización robusta y *minimax* del arrepentimiento; (ii) incorporar correlaciones entre parámetros inciertos con conjuntos basado en datos mediante el uso de PCA; y (iii) diseñar paquetes de políticas robustos sobre múltiples métricas con exploración de escenarios, selección iterativa e interpretabilidad.

El mensaje principal para las decisiones de inversión es pragmático: usar modelos como laboratorio de *stress test*, identificar inversiones habilitadoras y contingentes, y diseñar políticas monitorizables que reduzcan exposición a resultados extremos y generen un clima favorable a las inversiones necesarias. En lugar de preguntar “¿qué va a pasar?”, la pregunta útil es “¿qué decisiones siguen siendo aceptables si el futuro no se parece a lo esperado?”.

Quedan retos abiertos: incorporar decisiones secuenciales con información imperfecta de forma más explícita, ampliar la representación de costes de implementación de políticas, integrar participación de actores para definir pesos y umbrales de robustez, y conectar de

manera más directa con modelos de mercado y finanzas. Sin embargo, el marco desarrollado muestra que es posible avanzar hacia una planificación más realista y útil para la inversión, manteniendo tratabilidad y transparencia.

Desde una perspectiva académica y aplicada, se identifican cuatro líneas prometedoras:

- Decisión secuencial explícita: incorporar de manera más directa decisiones por etapas con aprendizaje, incluyendo reglas adaptativas endógenas.
- Costes de implementación y restricciones políticas: integrar mejor los costes administrativos, la aceptación social y las restricciones institucionales, para conectar la robustez técnica con la viabilidad política.
- Vínculo con modelos de mercado y con riesgos financieros: traducir los resultados de robustez en métricas relevantes para la financiación (por ejemplo, el riesgo de activos varados).
- Participación y deliberación: usar herramientas como openMASTER en procesos participativos, donde los pesos de métricas y umbrales se definan con actores, aumentando su legitimidad.

El desarrollo de modelos y métodos no sustituye la toma de decisión; la complementa. En la transición energética, la calidad del proceso de decisión es, en sí misma, una inversión institucional.

Finalmente, conviene subrayar una dimensión ética y de calidad democrática: las decisiones de inversión en transición afectan a la distribución de costes y beneficios y, por tanto, requieren deliberación pública informada. Los modelos *open source* y las metodologías de robustez pueden contribuir a ese objetivo si se usan con transparencia: documentando supuestos, mostrando rangos de resultados, explicitando *trade offs* y evitando presentar una trayectoria “óptima” como inevitable. En ese sentido, el mayor aporte de estas herramientas puede ser precisamente cambiar la conversación: de debatir sobre una trayectoria “óptima” presentada como inevitable a debatir sobre supuestos, vulnerabilidades y opciones de adaptación explícitas.

REFERENCIAS

BERTSIMAS, D., & SIM, M. (2004). The Price of Robustness. *Operations Research*, 52(1). <https://doi.org/https://doi.org/10.1287/opre.1030.0065>

HAASNOOT, M., KWAKKEL, J. H., WALKER, W. E., & TER MAAT, J. (2013). Dynamic adaptive policy pathways: A method for crafting robust decisions for a deeply uncertain world. *Global Environmental Change*, 23(2), 485-498. <https://doi.org/10.1016/j.gloenvcha.2012.12.006>

MARCHAU, V. A. W. J., WALKER, W. E., BLOEMEN, P. J. T. M., & POPPER, S. W. (Eds.). (2019). *Decision Making under Deep Uncertainty: From Theory to Practice*. Springer International Publishing. <https://doi.org/10.1007/978-3-030-05252-2>

PFENNINGER, S., HAWKES, A., & KEIRSTEAD, J. (2014). Energy systems modeling for twenty-first century energy challenges. *Renewable and Sustainable Energy Reviews*, 33, 74-86. <https://doi.org/10.1016/j.rser.2014.02.003>

RODRIGUEZ-MATAS, A. F., LINARES, P., PEREZ-BRAVO, M., & ROMERO, J. C. (2024). Improving robustness in strategic energy planning: A novel decision support method to deal with epistemic uncertainties. *Energy*, 292, 130463. <https://doi.org/10.1016/j.energy.2024.130463>

RODRIGUEZ-MATAS, A. F., PEREZ-BRAVO, M., LINARES, P., & ROMERO, J. C. (2024). openMASTER: The open source Model for the Analysis of Sustainable Energy Roadmaps. *Energy Strategy Reviews*, 54, 101456. <https://doi.org/10.1016/j.esr.2024.101456>

RODRIGUEZ-MATAS, A. F., RUIZ, C., LINARES, P., & PEREZ-BRAVO, M. (2025). How energy strategies are shaped by the correlation of uncertainties. *Applied Energy*, 382, 125257. <https://doi.org/10.1016/j.apenergy.2024.125257>

RODRIGUEZ-MATAS, A. F., WILD, T. B., LINARES, P., LAMONTAGNE, J. R., & DOMINGUEZ-BARBERO, C. (2026). Designing robust energy policy packages under deep uncertainty: A multi-metric decision support framework. *Energy Policy*, 210, 115008. <https://doi.org/10.1016/j.enpol.2025.115008>

SAVAGE, L. J. (1951). The Theory of Statistical Decision. *Journal of the American Statistical Association*, 46(253), 55-67. <https://doi.org/10.1080/01621459.1951.10500768>

TRUTNEVYTE, E. (2016). Does cost optimization approximate the real-world energy transition? *Energy*, 106, 182-193. <https://doi.org/10.1016/j.energy.2016.03.038>

Información Publicaciones / Publications Information:

Funcas
Caballero de Gracia, 28
28013 Madrid
España / Spain
Tfno. / Phone: +34 91 596 54 81
publica@funcas.es

P.V.P.: Suscripción anual papel, 25 € (IVA incluido)
Edición digital, gratuita



ISSN: 2445-2726
Depósito Legal: M-7537-2016