TESIS DE LA UNIVERSIDAD DE ZARAGOZA

2021

284

Guillermo Badía Fraile

Financial performance of socially responsible investments: international evidence from alternative perspectives

Director/es

Torres Pradas, Lourdes

Ferruz Agudo, Luis

ISSN 2254-7606

Prensas de la Universidad Universidad Zaragoza

ISSN 2254-7606

Universidad Zaragoza
1542

Tesis Doctoral

FINANCIAL PERFORMANCE OF SOCIALLY RESPONSIBLE INVESTMENTS: INTERNATIONAL EVIDENCE FROM ALTERNATIVE PERSPECTIVES

Autor

Guillermo Badía Fraile

Director/es

Torres Pradas, Lourdes
Ferruz Agudo, Luis

UNIVERSIDAD DE ZARAGOZA
Escuela de Doctorado

2019

Doctoral Thesis

Financial performance of socially responsible investments: international evidence from alternative perspectives

PhD Candidate: Guillermo Badía

PhD Supervisors: Dr. Luis Ferruz and Dra. Lourdes Torres

- 2019 -

Department of Accounting and Finances
Faculty of Economics and Business
University of Zaragoza

Universidad Zaragoza

Facultad de Economía y Empresa Universidad Zaragoza

Index

Tables Index

Figures Index

General introduction

The main objective of this Doctoral Thesis is to evaluate the financial performance of socially responsible investments (SRI).

In recent decades, investment management has undergone a progressive adaption process in which conventional financial objectives are increasingly being complemented by non-financial attributes such as environment, social and governance (ESG) criteria. This trend reflects an increasing awareness of environmental, social, and ethical issues that is strongly influencing the purchase decisions of investors (Mollet and Ziegler, 2014). SRI appeals to investors who wish to go beyond the financial utility of their investments and also derive non-financial utility from holding securities that reflect their social values (Auer, 2016; Auer and Schuhmacher, 2016). Additionally, ESG issues are becoming an important part of investors' decision-making process by helping them to identify firms' long-term opportunities and risks. According to the 2016 Global Sustainable Investment Review, in 2016 there were $22.89 trillion of assets being professionally managed under responsible investment strategies globally, representing an increase of 25% since 2014. In 2016, 53% of managers in Europe used responsible investment strategies, this proportion being 22% in the US and 51% in Australia/New Zealand. This tendency has been ratified recently for the last two years. US asset managers considered ESG criteria across $11.6 trillion in assets, up 44 percent from $8.1 trillion in 2016 (USSIF, 2018). The EUROSIF (2018) report discloses sustained growth for most sustainable and responsible investment strategies. The past two years (2016-2018) show manifest signs of SRI becoming integral to European fund management.

The basic idea of SRI is to apply a set of screens to the available investment universe, in order to select or exclude assets based on ESG criteria (Auer, 2016). In practice, there is a range of SRI strategies, such as integration, positive/best-in-class screening, ethical/negative screening, governance and engagement, etc. All of these aim to drive funds towards socially responsible firms with constructive sustainable projects and policies. From an investors' perspective, the critical issue is whether socially responsible stock selection leads to gains or losses in terms of financial performance. On the firms' side, the question is whether spending resources on corporate social responsibility (CSR) practices will render benefits for the firm and increase its value. If doing good is indeed linked to doing well, firms may be led to behave in a more

sustainable way. A positive relationship between social and financial performance would even legitimize CSR on economic grounds (Margolis et al. 2009).

The growth of SRI and its consequences has stimulated empirical studies assessing financial behaviours. An important stream of the literature has focused on the financial performance of SRI mutual funds. In general, these studies find that there are no significant differences between the performance of SRI mutual funds and conventional funds (Leite et al. 2018).[1] However, assessing the financial impact of SRI by evaluating the performance of actively managed SRI mutual funds has some shortcomings. For instance, as Brammer et al. (2006), and Kempf and Osthoff (2007) point out, there are confounding effects - such as fund manager skills and management fees - that may make it difficult to identify the performance that is due to the social characteristics of the underlying holdings. Furthermore, the evidence of Utz and Wimmer (2014), Humphrey et al. (2016), and Statman and Glushkov (2016) suggests that the 'socially responsible' label may be more akin to a marketing strategy, thus raising doubts among investors that an SRI fund is really socially responsible. As a consequence, investors may find it difficult to know the extent to which an SRI fund is really considering social criteria in its selection process. To overcome the limitations associated to studies on actively managed SRI mutual funds, an alternative approach to evaluate the financial effects of SRI involves evaluating the performance of synthetic portfolios formed on assets' social characteristics. In this Doctoral Thesis, we follow this approach to evaluating socially responsible investments.

This Doctoral Thesis is organized in two sections. The first includes chapters 1 and 2 in which we evaluate some methodological aspects related to a financial performance measure which is used to assess the financial performance of SRI in Section two. The second Section includes Chapters 3, 4, 5, and 6 in which we evaluate the financial performance of SRI from different perspectives.

First Section.

In Chapter 1 we assess the usefulness of a sector investment strategy based on the three-factor Fama and French (1992) model. We develop an investment process that is, as far as we know, new by including stocks that are undervalued with respect to their sector indices in a portfolio. We take as the relevant market factor the sector index to which

[1]For a review of studies on the performance of SRI equity funds see, for instance, Capelle-Blancard and Monjon (2012), and Revelli and Viviani (2015).

firms belong. We base the strategy on the difficulty entailed in effectively choosing the appropriate market portfolio (Roll, 1977).Our main objective in this chapter is to test whether it is possible to consistently achieve extra-financial returns by means of a sector strategy using the Fama and French model (1992) as a basis for decision-making.

In Chapter 2 we evaluate whether the Fama and French (1992) model may be adapted to become a more versatile and flexible tool, capable of incorporating variations of firms characteristics in a more dynamic form. We pay attention to the procedure that Fama and French (1992) follow to form the risk factors. They take annual data, and the value and size portfolios are assessed once a year, maintaining invariability during the whole period. However, we note that firms' characteristics can change during any given 12-month period. We argue that, over time, firms' valuation may change as a result of variations in its market price, size or book price, and we are aware that the Fama and French (1992) model does not accurately reflect these dynamics. Our main objective in this chapter is to test the effectiveness of the model by taking month-to-month data and reforming the value and size portfolios at the end of each month, aiming to develop a more dynamic and adaptable tool.

Second Section.

In Chapter 3 we evaluate the financial performance of portfolios that can be formed by socially conscious retail investors compared to conventional investments. We note that most previous studies evaluating the financial performance of SRI are conducted from the perspective of institutional investors' investment decisions and not from the perspective of retail investors who wish to hold SRI portfolios. Nonetheless, there has been a considerable increase in the popularity of SRI among retail investors (Benijts, 2010). Nilsson (2015) highlights that retail investors choose to devote at least part of their funds to investments that include some kind of social or environmental concerns, thereby having become an important factor in shaping SRI. According to the 2016 Global Sustainable Investment Review, although the SRI market in most of the regions is dominated by professional institutional investors, retail investors' interest in SRI is gaining relevance. Indeed, the relative proportion of retail SRI investments in Canada, Europe and the United States increased from 13 percent in 2014 to 26 percent at the start of 2016 (GSIA, 2016). Furthermore, over one third of SRI assets in the United States come from retail investors. The objective of this chapter is to assess the performance of portfolios that can be formed by socially responsible retail investors

compared to conventional investments. We use several financial performance measures. Among others, that developed in chapter 2 of this Doctoral Thesis. As a relevant point to retail investors, we use stocks listed on a source freely available to the public that any retail investor may access. Additionally, we analyse the impact of different market states on the financial performance of SRI portfolios. Recent research shows that the performance of SRI equity funds (Nofsinger and Varma, 2014; Becchetti et al., 2015, Leite and Cortez, 2015), SRI fixed-income funds (Henke, 2016), and socially responsible stocks (Brzeszczyński and McIntosh, 2014; Carvalho and Areal, 2016) is sensitive to different market states (e.g., expansion and recession periods).

In Chapter 4 we evaluate the financial performance of international stock portfolios based on CSR criteria. We note that previous studies that address the performance of socially screened synthetic portfolios suffer from some limitations and inconsistencies, namely, (1) the majority of prior evidence only refers to the US and European stock markets; (2) with the exception of Badía et al. (2017), previous studies do not compare the performance of SRI portfolios of different regions worldwide; (3) there are studies that measure CSR through one of its individual dimension only, whereas others consider an aggregate construct of CSR; (4) most studies do not evaluate the influence of specific industries on the financial performance of SRI stock portfolios; (5) in several studies assessing European firms, undersized samples are used; (6) up-to-date evidence is lacking; and (7) some researchers who split sample periods merely into sub-periods to evaluate a 'time effect', i.e., whether SRI returns were better in earlier years and yet declined in more recent periods, may have neglected an important effect, specifically, the impact of different market states. Our main objective in this chapter is to evaluate the financial performance of international stock portfolios based on CSR criteria aiming to overcome previous limitations in the evaluation of SRI stock portfolio performance. We form portfolios of stocks with high and low sustainability scores and investigate the performance of such portfolios using multi-factor models. In this chapter, we extend the analysis on the impact of including socially responsible screens on investment portfolios performance to additional geographical areas (North America, Europe, Japan, and Asia Pacific); we compare the financial performance of SRI portfolios of these regions to each other; we form portfolios based on an aggregate measure of CSR as well as on three of its specific ESG dimensions; we evaluate the influence of specific industries on the financial performance of SRI stock portfolios; and finally, we assess the financial

performance of SRI stock portfolios over different market states: bear, bull and mixed market periods.

In Chapter 5 we evaluate the financial performance of government bond portfolios formed according to ESG criteria. We note that although the concept of SRI was originally related to stock selection, the proportion of portfolio investors applying SRI criteria to bonds has grown significantly in recent years. According to the European Sustainable Investment Forum (EUROSIF, 2016), equities represented over 30% of SRI assets in December 2015, a significant decrease from the previous year's 50%. Meanwhile, there was a strong increase in bonds from the 40% registered in December 2013 to 64%. Both corporate bonds and government bonds underwent a remarkable growth. The former rose from 21.3% to 51.17% of the bond allocation, while the latter increased from 16.6% to 41.26%.In this regard, the financial implications of ESG screening processes on corporate bonds may be closely related to stock selections since corporate bonds are associated with firms. Indeed, previous studies (e.g., Derwall and Koedijk, 2009; Leite and Cortez, 2016) which evaluate the financial performance of mutual funds that invest in socially responsible fixed-income stocks, find that the average SRI bond funds performed similarly to conventional funds. These results are in line with most empirical studies about the performance of SRI funds, which show that they tend to have a similar performance to their conventional peers (Revelly and Viviani, 2015). However, ESG screening processes on government bonds, since they are not related to firms, can help gain an in-depth understanding of SRI consequences for alternative assets. Despite the SRI government bond market growth and the development of country ratings based on ESG factors in recent years, the link between government bond returns and country performance in terms of ESG concerns has been overlooked. In fact, to the best of our knowledge, no previous research has evaluated the financial performance of responsible government bond investments. The main objective of this chapter is to fill this gap. We assess the financial performance of government bond portfolios formed according to ESG criteria. We thus open a discussion on the financial performance of SRI for an alternative asset to firms.In contrast to previous studies, which apply firm sustainability ratings, we use sustainability ratings related to countries.

In Chapter 6 we ascertain a less assessed aspect in CSR: distinguishing between investments in material versus immaterial sustainability issues. We note that only firms focused on material sustainability issues associated with their main operations should

achieve a competitive advantage and obtain a higher social and financial performance. CSR activities and innovations should be performed on material aspects since otherwise a positive effect on financial performance is not expected. Indeed, investments on immaterial issues may involve additional corporate costs without a social and financial performance associated return. Focusing on material issues is important for firms since they do investments in social aspects that truly affect their operations. Despite issues as product safety, climate change, and resource intensity have impacts across several industries, as Hertz et al. (2016) note, those effects often vary to a great extent from one industry to the next. Risks may be everywhere, although they are indeed also particular. As a consequence, firms of specific industries have their particular sustainability profiles. Thus, a firm investing and reporting on material sustainability issues is likely achieved positive financial performance. Meanwhile, a firm investing on material but also on immaterial sustainability issues is likely not achieved superior financial performance. In this chapter, the main objective is to assess the financial performance of stock portfolios formed according to material and immaterial CSR issues. Khan et al. (2016) show that US firms with strong performance on material aspects outperform firms with poor performance on material topics. Our dataset includes companies from US and Europe. We thus extend the previous evidence of Khan et al. (2016) to European firms. Evaluating firms from US and Europe is particularly interesting given the heterogeneity in the patterns of development of SRI across countries (Neher and Hebb, 2015). Furthermore, we use firm' scores from an original dataset that integrates the SASB Materiality Map standards which, to our knowledge, has not been used before.

Introducción general

El objetivo principal de esta Tesis Doctoral es evaluar el desempeño financiero de las inversiones socialmente responsables (ISR).

En las últimas décadas, la gestión de inversiones ha experimentado un proceso progresivo de adaptación en el que los objetivos financieros convencionales se han complementado con atributos no financieros como los criterios medioambientales, sociales y de gobernanza (ESG). Esta tendencia refleja una creciente conciencia sobre cuestiones ambientales, sociales y éticas que influye de manera importante en las decisiones de compra de los inversores (Mollet y Ziegler, 2014). La ISR atrae a inversores que desean ir más allá de la utilidad financiera de sus inversiones y que esperan una utilidad no financiera que refleje sus valores sociales (Auer, 2016; Auer y Schuhmacher, 2016). Los aspectos ESG se están convirtiendo en una parte importante del proceso de toma de decisiones de los inversores al ayudarles a identificar oportunidades y riesgos en el largo plazo. De acuerdo con el Global Sustainable Investment Review de 2016, en 2016 hubo 22,89 billones de dólares gestionados profesionalmente en el marco de estrategias de inversión responsable a nivel mundial, lo que representa un aumento del 25% desde 2014. En 2016, el 53% de los gestores en Europa utilizaron estrategias de inversión responsable, siendo esta proporción del 22% en EE.UU. y del 51% en Australia/Nueva Zelanda. Esta tendencia se ha ratificado para los dos últimos años. Los gestores de activos estadounidenses consideraron criterios ESG en su gestión por valor de 11,6 billones de dólares, un 44 por ciento más que los 8,1 billones de dólares de 2016 (USSIF, 2018). El informe EUROSIF (2018) también revela un crecimiento sostenido en Europa de las estrategias de inversión sostenibles. Los dos últimos años (2016-2018) muestran signos manifiestos de que la ISR se está convirtiendo en parte integrante de la gestión de los fondos europeos.

La idea básica de la ISR es aplicar un conjunto de filtros al universo de inversión disponible con el fin de seleccionar o excluir activos en función de criterios ESG (Auer, 2016). En la práctica, existen diferentes estrategias ISR, como la integración, la selección positiva/*best-in-class*, la selección ética/negativa, la gobernanza, el compromiso, etc., todas ellas con el objetivo de dirigir los fondos hacia empresas socialmente responsables con proyectos y políticas constructivas y sostenibles. Desde la perspectiva de los inversores, la cuestión crítica es si la selección de acciones socialmente responsable conduce a ganancias o pérdidas en términos de rendimiento financiero. Por parte de las empresas, la cuestión es si el gasto de recursos en prácticas

de responsabilidad social de las empresas (RSE) redundará en beneficio de la empresa y aumentará su valor. Si hacer el bien (social y medioambiental) está vinculado a hacerlo bien (financieramente), las empresas podrían verse incentivadas a comportarse de manera más sostenible. Una relación positiva entre el desempeño social y el financiero legitimaría incluso la RSE sobre razones económicas (Margolis et al. 2009).

El crecimiento de la ISR y sus consecuencias ha estimulado la realización de estudios empíricos evaluando su comportamiento financiero. Una parte importante de la literatura se ha centrado en el rendimiento financiero de los fondos de inversión ISR. En general, estos estudios encuentran que no hay diferencias significativas en el desempeño financiero de fondos ISR y fondos de tipo convencional (Leite et al. 2018)[2]. Sin embargo, la evaluación del impacto financiero de la ISR mediante el análisis del rendimiento de los fondos de inversión ISR gestionados activamente presenta algunas deficiencias. Por ejemplo, como señalan Brammer et al (2006) y Kempf y Osthoff (2007), existen efectos confusos -como las habilidades de gestión del gestor y los honorarios y tasas por la gestión- que pueden dificultar la identificación del rendimiento de las ISR. Además, la evidencia de Utz y Wimmer (2014), Humphrey et al. (2016), y Statman y Glushkov (2016) sugiere que la etiqueta "socialmente responsable" puede ser una estrategia de marketing de los fondos, lo que suscitaría dudas entre los inversores sobre si un fondo ISR es realmente socialmente responsable. En consecuencia, los inversores pueden tener dificultades para saber en qué medida un fondo ISR tiene realmente en cuenta los criterios sociales en su proceso de selección. Para superar las limitaciones asociadas a los estudios sobre fondos de inversión ISR gestionados activamente, un enfoque alternativo para evaluar los efectos financieros de la ISR consiste en analizar el rendimiento de carteras sintéticas formadas utilizando características sociales, medioambientales y de gobernanza de las empresas. En esta Tesis Doctoral, seguimos este enfoque para evaluar las inversiones socialmente responsables.

Esta Tesis Doctoral está organizada en dos secciones. La primera incluye los capítulos 1 y 2 en los que se evalúan algunos aspectos metodológicos relacionados con una medida de rendimiento financiero que se utiliza para evaluar el rendimiento financiero de la ISR

[2]Para una revisión de los estudios sobre el rendimiento de los fondos de renta variable ISR, véase, por ejemplo, Capelle-Blancard y Monjon (2012), y Revelli y Viviani (2015).

en la sección dos. La segunda sección incluye los capítulos 3, 4, 5 y 6 en los que se evalúa el desempeño financiero de la ISR desde diferentes perspectivas.

Primera Sección.

En el Capítulo 1 se evalúa la utilidad de una estrategia de inversión sectorial basada en el modelo de tres factores de Fama y French (1992). En este capítulo desarrollamos un proceso de inversión, que hasta donde sabemos es nuevo, incluyendo en una cartera acciones que están infravaloradas con respecto a sus índices sectoriales, es decir, tomamos como factor de mercado relevante el índice sectorial al que pertenecen las empresas. Nuestro principal objetivo en este capítulo es comprobar si es posible conseguir de forma consistente una rentabilidad extraordinaria mediante una estrategia sectorial basada en el modelo de Fama y French (1992) para la toma de decisiones de inversión.

En el Capítulo 2 se evalúa si el modelo Fama y French (1992) puede convertirse en una herramienta más versátil y flexible, capaz de incorporar las variaciones en las características de las empresas de una forma más dinámica. Específicamente, prestamos atención al procedimiento que siguen Fama y French (1992) para formar los factores de riesgo. Ellos toman datos anuales y evalúan las carteras de valor y tamaño una vez al año, manteniéndolas invariables durante todo el período. Sin embargo, observamos que las características de las empresas pueden variar durante un periodo de 12 meses. Argumentamos que en ese periodo la valoración de una empresa puede cambiar como resultado de, por ejemplo, variaciones en su precio de mercado, su tamaño o su precio en libros; sin embargo el modelo de Fama y French (1992) no refleja con precisión esta dinámica. Nuestro principal objetivo en este capítulo es probar la eficacia del modelo tomando datos mensuales y reformando las carteras de valor y tamaño al final de cada mes para desarrollar una herramienta más dinámica y adaptable.

Segunda Sección.

En el Capítulo 3 se evalúa el rendimiento financiero de carteras que pueden formar inversores minoristas con conciencia social en comparación con inversiones convencionales. Observamos que la mayoría de los estudios previos que evalúan el rendimiento financiero de la ISR se llevan a cabo desde la perspectiva de las decisiones de inversión de los inversores institucionales y no desde la perspectiva de los inversores particulares que desean mantener carteras ISR. Sin embargo, ha habido un aumento

considerable de la popularidad de la ISR entre los inversores minoristas (Benijts, 2010). Nilsson (2015) destaca que los inversores particulares optan por dedicar al menos una parte de sus fondos a inversiones que incluyan algún tipo de preocupación social o medioambiental, convirtiéndose así en un factor importante en la configuración de la ISR. Según el Global Sustainable Investment Review de 2016, aunque el mercado ISR en la mayoría de las regiones está dominado por inversores institucionales profesionales, el interés de los inversores particulares por la ISR está adquiriendo relevancia. De hecho, la proporción relativa de inversiones en ISR al por menor en Canadá, Europa y Estados Unidos aumentó del 13 por ciento en 2014 al 26 por ciento a comienzos de 2016 (GSIA, 2016). El objetivo de este capítulo es evaluar el rendimiento de las carteras que pueden formar los inversores minoristas socialmente responsables en comparación con las inversiones convencionales. Utilizamos varias medidas de rendimiento financiero; entre otras, la desarrollada en el capítulo 2 de esta Tesis Doctoral. Como punto relevante para los inversores minoristas, para la selección de las empresas socialmente responsables acudimos a una fuente de información de acceso libre al público a la que puede acceder cualquier inversor minorista. Adicionalmente, en este capítulo analizamos el impacto que pueden tener diferentes estados del mercado (alcistas y bajistas) sobre el rendimiento financiero de las carteras ISR. Investigaciones recientes muestran que el rendimiento de fondos de renta variable ISR (Nofsinger y Varma, 2014; Becchetti et al., 2015, Leite y Cortez, 2015), fondos de renta fija de ISR (Henke, 2016) y empresas socialmente responsables (Brzeszczyński y McIntosh, 2014; Carvalho y Areal, 2016) son sensibles a diferentes estados del mercado.

En el Capítulo 4 evaluamos el desempeño financiero de carteras de acciones construidas con criterios de RSC a nivel internacional. Observamos que los estudios previos que abordan el desempeño de las carteras sintéticas socialmente responsables adolecen de algunas limitaciones e inconsistencias, a saber: (1) la mayoría de los estudios previos se centran en los mercados bursátiles de EE.UU. y Europa; (2) con la excepción de Badía et al. (2017), los estudios anteriores no comparan el desempeño de las carteras de ISR de diferentes regiones del mundo; (3) existen estudios que miden la RSC sólo a través de una de sus dimensiones individuales, mientras que otros consideran medidas agregadas de la RSC; (4) la mayoría de los estudios no evalúan la influencia de la industria en el desempeño financiero de las carteras de acciones ISR; (5) en varios de los estudios que evalúan a empresas europeas, se utilizan muestras de tamaño reducido; (6) falta evidencia actualizada; y (7) algunos investigadores simplemente dividen los

períodos de análisis en subperíodos para evaluar el "efecto de tiempo", sin embargo, es posible que se haya descuidado un efecto importante, el impacto de diferentes estados del mercado sobre el rendimiento financiero. Nuestro principal objetivo en este capítulo es evaluar el rendimiento financiero de carteras construidas sobre la base de criterios RSC superando las limitaciones previas. Formamos carteras de acciones con valoraciones de sostenibilidad altas y bajas e investigamos el rendimiento de dichas carteras utilizando modelos multifactoriales. En este capítulo, ampliamos el análisis sobre el impacto de la utilización de filtros socialmente responsables en el rendimiento de las carteras de inversión a otras áreas geográficas (Norteamérica, Europa, Japón y Asia-Pacífico); comparamos el rendimiento financiero de las carteras ISR de estas regiones entre sí; formamos carteras basadas en una medida agregada de RSE, así como en tres de sus dimensiones específicas ESG; evaluamos la influencia de la industria en el rendimiento financiero de las carteras de acciones ISR; y, por último, evaluamos el rendimiento financiero de las carteras de acciones ISR en diferentes estados de los mercados: alcistas, bajistas y períodos de mercados mixtos.

En el Capítulo 5 evaluamos el rendimiento financiero de carteras de deuda pública formadas según criterios ESG. Observamos que, aunque el concepto de ISR se relacionó originalmente con la selección de acciones, la proporción de inversores que aplican criterios ISR a bonos ha crecido significativamente en los últimos años. Según el Foro Europeo de Inversión Sostenible (EUROSIF, 2016), la renta variable representaba más del 30% de los activos de ISR en diciembre de 2015, lo que supone un descenso significativo respecto al 50% del año anterior. Por otra parte, se ha producido un fuerte aumento de los bonos, que han pasado del 40% registrado en diciembre de 2013 al 64%. Tanto los bonos corporativos como los bonos gubernamentales experimentaron un crecimiento notable. En este sentido, las implicaciones financieras de los procesos de selección ESG sobre bonos corporativos pueden estar estrechamente relacionadas con la selección de acciones, ya que los bonos corporativos están asociados a empresas. De hecho, estudios previos (por ejemplo, Derwall y Koedijk, 2009; Leite y Cortez, 2016) que evalúan el desempeño financiero de fondos que invierten en bonos de renta fija socialmente responsables, encuentran que en promedio tuvieron un desempeño similar al de los fondos convencionales. Estos resultados están en línea con la mayoría de los estudios empíricos sobre el desempeño de los fondos ISR que muestran que tienden a tener un desempeño similar al de sus pares convencionales (Revelly y Viviani, 2015). Sin embargo, los procesos de selección ESG sobre bonos gubernamentales, dado

que no están relacionados con las empresas, pueden ayudar a comprender las consecuencias de la ISR para activos alternativos. A pesar del crecimiento del mercado de deuda pública ISR y del desarrollo de calificaciones de los países basadas en factores ESG en los últimos años, se ha pasado por alto el vínculo entre la rentabilidad de la deuda pública y el rendimiento de los países en términos de preocupaciones ESG. De hecho, hasta donde sabemos, ninguna investigación previa ha evaluado el rendimiento financiero de las inversiones responsables en bonos gubernamentales. El objetivo principal de este capítulo es llenar este vacío. Evaluamos el rendimiento financiero de carteras de deuda pública formadas según criterios ESG. A diferencia de estudios previos, en los que se aplican calificaciones de sostenibilidad de las empresas, se utilizan calificaciones de sostenibilidad relacionadas con los países.

En el capítulo 6 se estudia un aspecto poco evaluado de la RSE: la distinción entre inversiones en cuestiones de sostenibilidad materiales e inmateriales. Sólo las empresas que se centran en cuestiones de sostenibilidad material asociadas a sus operaciones principales deberían lograr una ventaja competitiva y obtener un mayor rendimiento social y financiero. Las actividades de RSE y las innovaciones relacionadas deben realizarse sobre aspectos materiales, ya que de lo contrario no se debería espera un efecto positivo en el rendimiento financiero. De hecho, las inversiones en cuestiones inmateriales pueden implicar costes empresariales adicionales sin un rendimiento social y financiero asociado. Para las empresas es importante centrarse en las cuestiones materiales ya que de este modo invierten en aspectos sociales que realmente afectan a sus operaciones. A pesar de que temas como la seguridad de los productos, el cambio climático y la intensidad en el uso de los recursos tienen impactos en varias industrias, como señalan Herz and Rogers (2016), esos efectos a menudo varían en gran medida de una industria a otra. Los riesgos pueden estar en todas partes, aunque también son particulares. Como consecuencia, las empresas de diferentes industrias tienen sus perfiles de sostenibilidad particulares. Es probable que una empresa que invierte sobre temas de sostenibilidad material en su industria logre un desempeño financiero positivo. Mientras tanto, es probable que una empresa que invierte en cuestiones de sostenibilidad materiales pero también inmateriales no logre un rendimiento financiero superior. En este capítulo, el objetivo principal es evaluar el rendimiento financiero de carteras de acciones formadas en función de cuestiones de RSC materiales e inmateriales. Khan et al (2016) muestran que las empresas estadounidenses con un fuerte desempeño en aspectos materiales superan a las empresas con un desempeño

pobre en temas materiales. Nuestro conjunto de datos incluye empresas de estadounidenses y Europa. De este modo, ampliamos las pruebas anteriores de Khan et al. (2016) a las empresas europeas. La evaluación de las empresas estadounidenses y europeas es particularmente interesante dada la heterogeneidad de las pautas de desarrollo de la ISR en los distintos países (Neher y Hebb, 2015). En este capítulo utilizamos las puntuaciones de las empresas a partir de un conjunto de datos original que integra los estándares del Mapa de Materialidad SASB que, hasta donde sabemos, no se ha utilizado antes en este contexto.

First Section

Chapter 1: A Sector strategy from the Fama and French model

Luis Ferruz and Guillermo Badía

Faculty of Economics and Business, University of Zaragoza,
Zaragoza 50005, Spain.

Abstract

In this article, we test the degree of possible interest there may be in developing an investment strategy based on the three-factor Fama and French model (1992). To this end, we construct a sectorial strategy, taking as a market risk factor the sectorial index to which the securities belong. From our results, we conclude that the aforementioned strategy is of limited use, given that no extraordinary consistent yields are obtained. From this, we conclude that the hypothesis of the efficient market can be accepted.

Keywords Investment Strategies, Fama and French Model, Sectors, Market Efficiency

This study is published in the *Applied Economics Letters*, journal indexed in the Journal of Citation Report (JCR): Ferruz, L., & Badía, G. (2015). A sector strategy from the Fama and French model. *Applied Economics Letters*, *22*(18), 1511-1514.

1.1 Introduction

The objective of this work is to test whether it is possible to consistently achieve extraordinary yields by means of a sectorial strategy taking the three-factor Fama and French model (1992) as a basis for decision-making.

In our study, it is assumed that the risk factors identified by Fama and French (1992) are able to explain the behaviour of the securities yields, i.e. those that attain a level of profitability that does not match the predictions of the model and that are greater than the estimate, they could be considered undervalued, understanding that there exists an imbalance in the profitability-risk relationship and that it is possible to obtain extraordinary yields. A range of studies, such as Bartholdy and Peare (2005), Morana and Beltratti (2006) and Soumaré *et al*. (2013) have all dealt with the analysis and validation of the model.

Carhart (1997) identifies a fourth risk factor (Momentum) but it is excluded because some stocks of our sample are not traded all months and the return of the months in which the price does not change is zero. Denoting the momentum of month t by $MOM_{j,t} = R_{j,t-1} / R_{j,t-12}$, when the return of month t - 12 is zero, the value is not defined and missing value is generated.

The interest in analyzing each yield along with those of its sector is justified in the critiques of the complexity of knowing the true market portfolio (Roll, 1977). Taking advantage of this possible weakness for the development of an investment strategy, as something new, given what we know, we consider whether the securities are undervalued with regard to their sectorial indices.

If the objective is achieved, we will be capturing certain inefficiencies in the market to permit the attainment of extraordinary yields. Nevertheless, once the aforementioned strategy was widely known and implemented, the imbalance in prices would disappear, given the efficiency on which the model is based. On the contrary, by not achieving the objective, the hypothesis of market efficiency could be accepted.

The rest of the work is structured as follows. Section II presents the data and methodology used in the study. In Section III, our obtained results are discussed and we present our conclusions in Section IV.

1.2 Database and methodology

The period under analysis runs from January 2006 to December 2013, considering a total of 692 securities spread over eight sectors[3]. The contributions of companies, of sectorial indices and of the index to beat (FTSE 100) have been obtained from the Morningstar database. To be included in the sample, the securities must have remained in the UK market during the entire period of the study[4].

Given that the objective of the work is to contrast the efficiency of an investment strategy that takes a month as its period of reference for the composition-reconstruction of the portfolio, we use monthly prices and contributions to calculate profitability[5]. This profitability is calculated as the natural log of the quotient between the price at the close of the analyzed month, and the price at the close of the previous month.

Regarding the risk-free yield, the 3-month Treasury bill issued by the Bank of England is selected.

1.2.1 Obtaining the factors of the Fama and French model in each Sector

Due to the fact that the rebalancing of the portfolio is done month-to-month, the risk factors of the model are reconstructed for each sector in each month, taking monthly data. Thus, in the case of the Size factor, the stock market capitalization of the securities is taken at the end of each month, while the Book-to-Market (B/M) factor is obtained as a quotient between the price on the books of each security at the end of each month and its price at the end of each month in the open market.

1.2.2 Construction of the Portfolios of the Model in each Sector

The procedure for the construction of the portfolios SMB (Small minus Big) and HML (High minus Low) is the same as that followed by Fama and French (1992) with the difference being that these authors construct the portfolios once a year and maintain them during the whole period, and, in our study, given the objective of the analysis, the process is done month-to-month.

In this way, in the first place, the securities are arranged in each sector according to the value of the stock market capitalization. Within each sector, two equal groups are

[3]The sectors are described in the Appendix. The Financial sector is excluded, following Fama y French (1992).

[4]It is assumed, therefore, that some survival bias exists.

[5]Thus, liquidity problems derivative of the fact that certain securities in certain months have days when they are not traded is limited.

created by dividing the sample by the median, with the larger capitalization securities on one side (Big) and those of lesser capitalization (Small) on the other.

Subsequently, each of the groups is ordered from greater to lesser, according to the value of the ratio B/M, and is divided into three subgroups taking the values of the percentiles of 30 and of 70%. In this way, the value portfolio, the neutral portfolio and the growth portfolio of Big, as well as Small, companies are obtained in each sector (Table 1).

Table 1-1. Portfolios constructed in each sector to calculate the risk factors of the model

		Size	
	Portfolio	Small	Big
Ratio B/M	Low	Portfolio 1 (P1)	(P4)
	Medium	(P2)	(P5)
	High	(P3)	(P6)

Thus, SMB = [(P1+P2+P3)/3 – (P4+P5+P6)/3], whereas HML = [(P3+P6)/2 – (P1+P4)/2].

Given the objective of our study, we need to estimate the performance of the securities; hence the model is projected with this purpose for each one of the securities considered. Avramov and Chordia (2006) point out that using individual securities rather than portfolios avoids the loss of information when the securities are grouped together in portfolios and also avoids any bias in the data associated with the portfolio construction.

1.2.3 Beta estimations

Once the portfolios SMB and HML are obtained for each month and sector, time-series regressions are run in which the slopes of the model are estimated:

$$R_i - R_f = \alpha_i + \beta_i(R_S - R_f) + \beta_i SMB_S + \beta_i HML_S + \varepsilon_i$$

(Eq. 1-1)

Where R_i is the performance of the securities at a given moment and R_f is the return on the risk-free security. In this case (R_S – R_f) corresponds to the excess of the returns on the risk-free security from each of the sectorial indices to which each security belongs. SMB_S and HML_S are the average performances attained by the portfolios in each sector, α_i is the intercept term of the regression, β_i are the slopes of the model obtained by means of ordinary least squared, and ε_i is the disturbance term.

The period selected for the beta estimations is 5 years. This period, as Brooks (2014) maintains, is commonly used by other researchers. Nevertheless, Bartholdy and Peare (2005) carry out a study in which they determine that this is the optimum period for the estimation of these parameters.

1.2.4 Estimation of security performance

To continue, estimations of cross-section are carried out:

$$E(R_i) - R_f = \beta_{iM}\left[E(R_S) - R_f\right] + \beta_{iSMB}\,E(SMB_S) + \beta_{iHML}\,E(HML_S)$$

(Eq. 1-2)

Where $E(R_i)$ is the expected return of the stock and R_f is the return on the risk-free security. β_{iM}, β_{iSMB} and β_{iHML} are the betas obtained in the previous step, and $[E(R_S) - R_f)]$, $E(SMB_S)$ and $E(HML_S)$ are the averages of the risk factors in the previous 5 years.

We use the rolling window process by Fama and MacBeth (1973) to obtain the expected return of the stocks in each moment.

Estimating the performance of the securities, taking as a benchmark the sectorial index to which they belong, we compare it with the average performance actually obtained by the securities in that period and we add to the portfolio those securities whose profitability is greater than those estimated to be undervalued.

1.2.5 Performance of the developed strategy

Knowing the securities in each sector that will be included in the portfolio, the average profitability associated with it is calculated and compared to that obtained by the benchmark that aims to beat, the FTSE 100.

In addition to analyzing the excess profitability of the portfolio, the robustness of the strategy is verified by the statistical |Z| of Malkiel (1995). If the statistical value is above 1.96, it is considered that the results obtained are due to the construction of a strategy that allows consistently beating the market; otherwise, it is understood that the value may be due to mere chance. This indicator is used and analyzed by authors such as Reinker and Tower (2004).

1.3 Results

In Table 2, the results obtained are presented. As can be seen, by using the Fama and French model (1992) and applying it to different sectors, it is possible to beat the market

practically 53% of the time. The yields are somewhat higher than those obtained by the market in 2 of the 3 years in which the effectiveness of the strategy is analyzed and the values of the standard deviation are, similarly, lower for 2 of the 3 years considered. Nevertheless, since the total number of months in which the objective is met is 19, the statistical value of |Z| does not permit the assurance that the application of the developed strategy will consistently result in extraordinary profitability.

Table 1-2. Results of the sectorial strategy

Months analyzed			36	
Total months that beat the market			19	52.78%
		Z -test	0.118	<1.96
Total months that beat the market each year				
2011	6	50.00%		
2012	7	58.33%		
2013	6	50.00%		
Annual Performance	Portfolio	FTSE 100	Excess	Extraordinary Profitability?
2011	-0.0102	-0.0048	-0.0054	No
2012	0.0055	0.0047	0.0007	Yes
2013	0.0148	0.0112	0.0036	Yes
SD Annual	Portfolio	FTSE 100	Difference	Lowest SD?
2011	0.0303	0.0369	-0.0066	Yes
2012	0.0321	0.0288	0.0033	No
2013	0.0206	0.0336	-0.0130	Yes

1.4 Conclusions

In this work, we have attempted to ascertain the degree of interest that there may be in developing a strategy based on the three factors of the Fama and French model (1992) in making investment decisions aimed at obtaining extraordinary profitability. To this end, a sectorial strategy has been produced and, in view of the results, we can conclude that this strategy has reduced utility since it does not achieve extraordinary performance in a systematic and consistent manner. From this conclusion, we can determine that the hypothesis of the efficient market can be accepted.

References

Avramov, D. and Chordia, T. (2006) Asset pricing models and financial market anomalies, *Review of Financial Studies*, *19*, 1001-1040.

Bartholdy, J. and Peare, P. (2005) Estimation of expected return: CAPM vs. Fama and French, *International Review of Financial Analysis*, *14*, 407-427.

Brooks, C. (2008) *Introductory Econometrics for Finance*, Cambridge University Press.

Carhart, M. M. (1997) On persistence in mutual fund performance, *The Journal of Finance*, *52*, 57-82.

Fama, E. F. and French, K. R. (1992) The cross-section of expected stock returns, *The Journal of Finance*, *47*, 427-465.

Fama, E. F. and MacBeth, J. D. (1973) Risk, return, and equilibrium: empirical tests, *The Journal of Political Economy*, *81*, 607-636.

Malkiel, B. G. (1995) Returns from investing in equity mutual funds 1971 to 1991, *The Journal of Finance*, *50*, 549-572.

Morana, C. and Beltratti, A. (2006) Structural breaks and common factors in the volatility of the Fama–French factor portfolios, *Applied Financial Economics*, *16*, 1059-1073.

Reinker, K.S. and Tower, E. (2004) Index Fundamentalism Revisited, *The Journal of Portfolio Management*, *30*, 37-50.

Roll, R. (1977) A critique of the asset pricing theory's tests. Part I: on past and potential testability of the theory, *Journal of Financial Economics*, *4*, 129-176.

Soumaré, I., Aménounvé, E. K., Diop, O., Méité, D. and N'sougan, Y. D. (2013) Applying the CAPM and the Fama–French models to the BRVM stock market, *Applied Financial Economics*, *23*, 275-285.

Appendix A. Firms by sector

Table 1-3. Ap. A. Firms by sector

Sectors	No. of Securities
Basic Material	106
Consumer	192
Energy	62
Healthcare	48
Industrial	156
Technology	100
Telecommunications	16
Utilities	12

Chapter 2: Adapting and testing the Fama and French model, with some variations of company characteristics

Luis Ferruz and Guillermo Badía

Faculty of Economics and Business, University of Zaragoza,
Zaragoza 50005, Spain.

Abstract

We examine whether the Fama & French (1992) (F&F) model can be adapted to become a more versatile and flexible tool, capable of incorporating variations of company characteristics in a more dynamic form. For this, the risk factors are reconstructed at the end of each reading of monthly data. We argue that, over time, the evaluation of a company may change as a result of variations in its market price, size or book price, and we are aware that the F&F model does not accurately reflect these dynamics. Our results show that the adapted model is able to capture the behavior of a greater number of stocks than the original F&F model and risk factors are more significant when building them through our procedure. In addition, we carry out these adaptations during a period of instability in financial markets.

Keywords Financial models, Fama and French model, Risk factors

This study is published in the *Applied Economics Letters,* journal indexed in the Journal of Citation Report (JCR): Ferruz, L., & Badía, G. (2017). Adapting and testing the Fama and French model, with some variations of company characteristics. *Applied Economics Letters*, *24*(5), 342-345.

2.1 Introduction

In this article, we pay attention to the procedure that Fama and French (1992) (hereafter referred to as F&F) follow to build risk factors. F&F construct their risk factors from data taken annually, and the value and size of portfolios are assessed once a year, maintaining invariability during the whole period. However, it should be noted that variations can occur in the characteristics of a company during any given 12-month period, which will not be accounted for by the F&F procedure.

The F&F model has been widely used and analyzed in the literature recently (Abhakorn et al.2013; Eraslan 2013; Gregory et al. 2013; Soumaré et al. 2013; Nichol and Dowling 2014; Zhong et al. 2014; and Ferruz and Badía 2015, among many others). Our main research goal is to test the capacity of the model, taking month-to-month data and rebuilding the value and size portfolios at the end of each month, with the aim of developing a more dynamic and adaptable tool.

This approach has two clear implications. First, the Book-to-Market (BM) ratio varies according to the characteristics of the company at any given moment. Although we can expect that the numerator or book price appears invariant for an entire period, the denominator or market price does vary and, therefore, a company may be moving between different value portfolios during the year without being captured, i.e. between low, medium and high portfolios. Second, our approach provides greater variability in the size factor, which allows us to capture variations in the capitalization of the company as a result of, among other things, price fluctuations of the stock, which can have consequences for the classification of a company (such as, Big or Small).

The variation that may arise in the characteristics of a company, causing it to move between different value and size portfolios, has a direct impact on the associated return of the portfolios. Thus, our approach allows us to form new portfolios every month by utilising the true set of characteristics, and the yields reflected by these portfolios are better suited to the situation at any given moment.

To test the ability of the adapted model, the significance of the model as a whole and the individual coefficients considered in the regression are checked.

The rest of the article is organized as follows. In Section II, we present the data and methodology, in Section III, we offer the results of the empirical analysis and in Section IV, we discuss our conclusions.

2.2 Data and methodology

Our data covers the period from January 2006 to December 2010. Five-year intervals to estimate the coefficients, as argued by Brooks (2008), are often used for this purpose, and the financial crisis period is included in order to test the effectiveness of our recursive construction process of the factors, in a highly-volatile environment during which stock markets suffered major shocks, as did the valuations of companies and their variables.

We sample a total of 692 nonfinancial firms trading in the UK[6]Market. Company monthly prices and the rest of the necessary data are taken from the Morningstar Database. Returns are calculated as the natural logarithm of the quotient between the price at time *t* and price at *t-1*. The FTSE All Share Index is used as a proxy for the market portfolio, and the 3-month UK Treasury Bill rate is the risk-free asset.

We also use risk factors built by F&F in order to compare the results of our proposal. These data are obtained from the website of Kenneth French.

To build the adapted model, we reconstruct the risk factors at the end of each month from the monthly data. Thus, to obtain the size factor, we take the stock exchange capitalization at the end of each month, and the BM factor is calculated as the quotient between the book price and the share market price, both also taken at the end of every month.

Subsequently, we proceed to construct the SMB (Small minus Big) and HML (High minus Low) portfolios. The procedure is the same as that followed by F&F, except that they build their portfolios annually (in June) and hold them during the entire period, whereas we conduct our procedures on a month-to-month basis.

Thus, to construct the SMB portfolio, we rank the securities by capitalization value at the end of each month and establish two groups, dividing the sample by the median. In this way, we have the large-capitalization assets on one side and the small-capitalization assets on the other.

Thereafter, every group is ranked from highest to lowest according to the BM value and divided into three subgroups, taking the same values as F&F, with the percentiles of 30% and 70%. In this way, the Value portfolio (High), the Neutral portfolio (Medium) and the Growth portfolio (Low) of Big and Small companies are obtained.

[6]Financial companies are excluded, for the same reasons as Fama and French (1992).

The SMB risk factor is the average return associated with the difference between the average return portfolio of small-cap companies and the mean return portfolio of large-cap companies, whereas the HML factor risk is the average return portfolio of value assets minus the average return portfolio of growth assets.

In this way, we create six portfolios, by size and BM ratio, and reconstructed each month according to changes in the characteristics of the companies, with the target being to provide greater reaction and adaptability to the model under certain contingencies.

Once the portfolios SMB and HML are obtained for each month, regressions are run for each security in which the coefficients of the model are estimated.

$$R_i - R_f = \alpha_i + \beta_i RMRF + \beta_i SMB + \beta_i HML + \varepsilon_i$$

(Eq. 2-1)

Where R_i is the performance of the securities, R_f is the return on the risk-free asset, *RMRF* is the Market risk factor (i.e. the excess return of the Benchmark on the risk free asset), *SMB* is the difference between the Small-cap stocks portfolio returns and the Large-cap stocks portfolio returns and *HML* is the difference between the High securities portfolio returns and the Low securities portfolio returns;α_i is the intercept term of the regression, β_i is the slope of the model and ε_i is the disturbance term.

2.3 Empirical analysis

The estimation results for each model are presented in Table 1, which allows us to observe and to compare the ability of the factors, and the model as a whole, according to both approaches. For each coefficient, the number of times that it appears significant is counted, and also its significance level, in each one of the 692 regressions of each model. Each percentage is calculated on the total titles. The accumulated value is the total to 10%, i.e. the sum of 1%, 5% and 10%.

Taking the 1% significance level, it can be seen that the RMRF factor on the F&F model appears significant to 56 regressions (8.09%), while on the adapted model, it is shown to be significant to 471 regressions (68.06%). In the case of the HML factor, it appears significant to 102 regressions (14.74%) of the adapted model and only to 3 regressions (0.43%) on the F&F model. The SMB factor is shown to be significant in 99 titles (14.31%) in our adapted model, while in the F&F model it appears in 28 titles (4.05%).

When we observe the cumulative total, i.e. increasing the significance level up to 10%, we appreciate that both RMRF and HML factors appear on more significant occasions according to the adapted model (74.86% and 33.38%, respectively) than with the F&F model (34.39% and 11.71%, respectively).For the SMB factor, this appears significant more often for the F&F model, but we note that, in most cases (22.25%), it occurs at the 10% level.

As for intercept values, the case of the F&F model appears significant at 1% for all regressions (100%), while for the adapted model it is only in 12 regressions (1.73%).A good model specification produces intercepts that are indistinguishable from zero (Merton, 1973). As F&F indicate, the intercept estimation provides a simple measure, and in turn a formal test of how the different factors capture the average performance. Therefore, these results show that most of the stock behavior is captured by the risk factors considered in the Adapted model.

The F-Statistic values settle results. When we consider the 1% level, the F&F model fits the behavior of 16.04% titles, while the adapted model captures 96.82%, a more than relevant amount. When we broaden the significance level and observe the accumulation, the adapted model captures almost the entire stock behavior (99.28%), while the F&F model captures just over half (53.03%).

Table 2-1. Significance results of the coefficients on F&F model and on adapted model

F&F model	Significance Level							
	1%	%	5%	%	10%	%	Accumulated	%
Intercept	692	100	0	0.00	0	0.00	692	100
RMRF	56	8.09	100	14.45	82	11.85	238	34.39
HML	3	0.43	32	4.62	46	6.65	81	11.71
SMB	28	4.05	91	13.15	154	22.25	273	39.45
F-Statistic	111	16.04	162	23.41	94	13.58	367	53.03
Adapted model	Significance Level							
	1%	%	5%	%	10%	%	Accumulated	%
Intercept	12	1.73	31	4.48	51	7.37	94	14
RMRF	471	68.06	20	2.89	27	3.90	518	74.86
HML	102	14.74	79	11.42	50	7.23	231	33.38
SMB	99	14.31	90	13.01	71	10.26	260	37.57
F-Statistic	670	96.82	13	1.88	4	0.58	687	99.28

For each coefficient, the number of times that it appears significant is counted, along with its significance level, in each one of the 692 regressions with each model. Percentages are presented in italics and each percentage is calculated on the total titles. The accumulation is the total to 10%, i.e. the sum of 1%, 5% and 10%.

These results are presented as being of particular relevance to investors and Management, not only because the ability of the adapted model to fit the returns of companies is better than the F&F model, but also because it is carried out during a period of high instability. Often, the effectiveness of this type of model is criticized when market conditions are in crisis and the efficient market hypothesis is compromised.

2.4 Conclusions

The results of our proposed adaptation of the F&F model indicate that it can be converted into a more flexible, versatile and dynamic tool, since the construction of the risk factors taking monthly data allows us to adapt them more recurrently, resulting in an improved ability to capture the variations that may arise in the characteristics of companies in the course of any given period.

It is important to stress that the intercept and F-Statistic values show that the adapted model is able to fit the behavior of almost all stocks, while the F&F model only does so for just over half.

Thus, in view of the results, we maintain that this procedure has important implications, as well as presenting a more efficient model than the original F&F model, with its efficiency proven when applied to a period of crisis.

References

Abhakorn, P., Smith, P. N. and Wickens, M. R. (2013) What do the Fama–French factors add to C-CAPM? *Journal of Empirical Finance, 22*, 113-127.

Brooks, C. (2008) *Introductory econometrics for finance*, Cambridge University Press.

Eraslan, V. (2013) Fama and French three-factor model: evidence from Istanbul stock exchange, *Business and Economics Research Journal, 4*, 11-22.

Fama, E. F. and French, K. (1992) The cross-section of expected stock returns, *The Journal of Finance, 47*, 427–465.

Ferruz, L. and Badía, G. (2015) A sector strategy from the Fama and French model, *Applied Economics Letters, 22*, 1511–1514.

Gregory, A., Tharyan, R. and Christidis, A. (2013) Constructing and testing alternative versions of the Fama–French and Carhart models in the UK, *Journal of Business Finance and Accounting, 40*, 172-214.

Merton, R. C. (1973) An intertemporal capital asset pricing model, *Econometrica, 41*, 867-887.

Nichol, E. and Dowling, M. (2014) Profitability and Investment Factors for UK Asset Pricing Models, *Economics Letters, 125*, 364-366.

Soumaré, I., Aménounvé, E. K., Diop, O., Méité, D. and N'sougan, Y. D. (2013) Applying the CAPM and the Fama–French models to the BRVM stock market, *Applied Financial Economics, 23*, 275-285.

Zhong, A., Limkriangkrai, M. and Gray, P. (2014) Anomalies, risk adjustment and seasonality: Australian evidence, *International Review of Financial Analysis, 35*, 207-218.

Second Section

Chapter 3: The performance of social responsible investing from retail investors' perspective: international evidence

Guillermo Badía[1], Luis Ferruz[1], and María Ceu Cortez[2]

[1]Faculty of Economics and Business, University of Zaragoza,
Zaragoza 50005, Spain.

[2]NIPE - School of Economics and Management, University of Minho,
Gualtar, 4710-057, Braga, Portugal.

Abstract

This paper investigates the performance of socially responsible investment (SRI) portfolios compared to conventional investments. Adopting a retail investor's perspective, we provide evidence of SRI financial performance at the worldwide level as well as at the regional level, for five regions (America, Europe except UK, UK, Pacific region and Emerging markets). Furthermore, we analyse the impact of different market states on the financial performance. Our results show that over the period 2005 to 2014, SRI portfolios outperform conventional investments. During bear market periods, the financial performance is neutral for both portfolios, whereas during bull market periods SRI portfolios statistically outperform the conventional portfolio. This outperformance is related to a positive and statistically significant exposure to the size and value risk factors. At the regional level, the results show statistical differences in the financial performance of regional SRI portfolios. These results suggest that country-specific factors may affect the relationship between corporate social and financial performance.

Keywords Socially responsible investing; Retail investors; Portfolio performance evaluation; Market states; International portfolios

This study was awarded in the XIX AECA International Conference for its scientific quality and its contributions. 27th-29th September, Santiago de Compostela, Spain.
Currently, this study is under review in the International Journal of Finance and Economics, journal indexed in the Journal of Citation Report (JCR).

3.1 Introduction

The interest in socially responsible investing has increased significantly over the last decades in both academic research (Scholtens, 2015) and investors' practices (Ferruz et al., 2012; van Duuren et al., 2016). Investors are increasingly willing to incorporate into their investment decisions not only financial criteria (returns and risk), but also the non-financial attributes of SRI (Benson and Humphrey 2008; Nicolosi et al. 2014). The worldwide growth of SRI, as Nilsson (2015) notes, is taking place despite some scepticism on its effects, such as a limited set of SRI investment options and loss of portfolio diversification. Nevertheless, proponents of SRI claim that socially screened investments may result in a higher financial performance. This argument is supported by many empirical studies that document a positive link between corporate social responsibility and corporate financial performance and valuation.[7] Additionally, the issue of whether considering social screens has a positive or negative effect on the financial performance of investment portfolios has been the focus of many empirical studies (e.g., Kempf and Ostoff, 2007; Statman and Glushkov, 2009; Renneboog et al., 2008; Borgers et al., 2013). As the review studies of Capelle-Blancard and Monjon (2012) and Revelli and Viviani (2015) point out, most studies do not find statistically significant differences between the financial performance of SRI portfolios compared to conventional portfolios.

However, we note that most previous studies evaluating the financial performance of SRI are conducted from the perspective of institutional investors' investment decisions and not from the perspective of retail investors who wish to hold SRI portfolios. Nonetheless, there has been a considerable increase in the popularity of SRI among retail investors (Benijts, 2010). Nilsson (2015) highlights that retail investors choose to devote at least part of their funds to investments that include some kind of social or environmental concerns, thereby having become an important factor in shaping SRI. According to the 2016 Global Sustainable Investment Review, although the SRI market in most of the regions is dominated by professional institutional investors, retail investors' interest in SRI is gaining relevance. Indeed, the relative proportion of retail SRI investments in Canada, Europe and the United States increased from 13 percent in 2014 to 26 percent at the start of 2016 (GSIA, 2016). Furthermore, over one third of

[7]For a more in-depth discussion of the empirical studies in the field, see for example, the review studies of Margolis and Walsh (2003), Orlitzky et al. (2003), Margolis et al. (2009), Lu et al. (2014) and Javed et al. (2016).

SRI assets in the United States come from retail investors. There are at least two issues that are relevant for retail investors who wish to choose socially responsible investments. First, many studies in the field use data on stocks' social scores from proprietary and expensive databases. We argue that access to information sources is more limited and restricted for retail investors than for institutional investors. Retail investors have little choice but to use open sources of information that are freely available, whereas institutional investors have access to expensive information sources and databases. There are several studies that consider the perspective of retail investors in following SRI strategies based on free and accessible information on corporate social performance, but we note that they are mainly focused on the US and UK equity markets (e.g., Brammer et al., 2009; Edmans, 2011; Filbeck, 2013; Brzeszczyński and McIntosh, 2014; Carvalho and Areal, 2016).

Second, as emphasised by Osthoff (2015), most studies on SRI focus on financial products such as mutual funds. (e.g., Bauer et al., 2005, Statman, 2006; Schröder, 2007; Renneboog et al., 2008; Cortez et al., 2009, 2012). Retail investors may in fact be interested in investing in actively managed SRI mutual funds. However, as Auer and Schumacher (2016) point out, evaluating the impact of incorporating social screens by analysing the performance of mutual funds has some limitations. A major problem is that there is some evidence that the label 'socially responsible' might be more of a marketing strategy, thus not assuring investors that a SRI fund is truly socially responsible. The issue of whether SRI funds are simply conventional funds in disguise has been recently debated in the literature. For instance, Wimmer (2013) shows that the social level of SRI funds largely disappears after two years. In turn, Utz and Wimmer (2014) find that that, on average, SRI funds do not hold more ethical stocks than conventional funds and that a mutual fund being classified as SRI does not ensure exclusion of socially controversial firms. The findings of Humphrey et al. (2016) reinforce the argument that SRI funds and conventional funds are not so different after all and Statman and Glushkov (2016) even find evidence of closet SRI funds, which are conventional funds that avoid investing in unethical stocks. In this context, retail investors may find it difficult to know the extent to which a SRI fund is really considering social criteria in their selection process. By constructing SRI portfolios, retail investors can be more confident that the companies that are included in their portfolios are indeed reflecting their social concerns. Furthermore, in countries where mutual funds are marketed by commercial banks, their interests may lead socially

responsible private investors towards products that are not suited for their social concerns. Banks are predominantly oriented to maximize profits and not the interests of depositors and investors. Graafland and Van de Ven (2011), for instance, document that during the credit crisis there were cases in which banks did not behave according to the moral standards they set for themselves, and claim that commercial practices and ethical values of financial professionals played a relevant role in the global financial crisis. Also, Van Hoorn (2015) points out that the financial services industry sometimes favours an environment that is highly conducive to unethical behaviour.

Considering that worldwide evidence regarding the possibility of SRI retail investors to yield positive financial performance is scarce as well as the problems socially conscious investors may face when trying to select true SRI funds, we focus on retail investors' perspective by forming portfolios based on social criteria. It is important to mention that currently, the technological developments in trading systems have reduced transaction costs and commissions, thereby encouraging retail investors to trade and leading to an increase in the trading volume and liquidity (Butt and Virk, 2017).

The purpose of this study is to analyse the performance of portfolios that can be formed by socially conscious retail investors compared to conventional investments. Following Nilsson's (2015) concerns that socially responsible retail investors need easy-to-use tools on social information, we form portfolios based on the stocks listed on the Global-100 'Global-100 Most Sustainable Corporations in the World' list (Global-100, hereafter), which is freely available to the public. We therefore use information sources that any retail investor may access in order to set up an investment portfolio that follows socially responsible investment criteria.

We contribute to the literature by extending the portfolio performance evaluation of portfolios formed on the basis of free and available social information to a worldwide context. Although Brzeszczyński and McIntosh (2014) also use the Global-100 to identify socially responsible opportunities, their analysis is limited to UK stocks included in the list. We consider that a worldwide analysis is relevant considering that the patterns of development of SRI are not homogenous across countries (Neher and Hebb, 2015). Furthermore, Hörisch et al. (2015) indicate that country-specific factors tend to affect the relationship between corporate social and financial performance. Our analysis includes all SRI Global-100 stocks without restricting the investigation to any specific country. Additionally, we analyse the impact of different market states on the financial performance of SRI portfolios. Recent research shows that the performance of

SRI equity funds (Nofsinger and Varma, 2014; Becchetti et al., 2015, Leite and Cortez, 2015), SRI fixed-income funds (Henke, 2016), and socially responsible stocks (Brzeszczyński and McIntosh, 2014; Carvalho and Areal, 2016) is sensitive to different market states (e.g., expansion and recession periods). We use a conditional model that includes dummy variables, in line with Nofsinger and Varma (2014), and Leite and Cortez (2015), in order to capture changes in both risk and performance across different market states. Although this analysis is conducted from a retail investor perspective, nonetheless, of course, institutional investors can take into account the results and conclusions reached in this empirical study for constructing their SRI strategies.

The structure of the paper is as follows: Section 2 presents a short overview of the relevant literature. Section 3 describes the data and Section 4 presents the research methods used. Section 5 contains and details the empirical results and Section 6 summarizes our main findings and offers some concluding remarks.

3.2 Prior literature

A stream of the SRI literature evaluates the performance of SRI portfolios formed on the basis of social scores provided by specialized rating agencies.[8] However, many of these studies use proprietary and expensive social databases which are not typically available to the general public. Studies that can be useful to retail investors, by forming SRI portfolios based on free and publicly available social information are scarce, and focus mainly on the US and the UK markets. For instance, Filbeck et al. (2009) analyse the performance of portfolios composed by the '100 Best Corporate Citizens' published by *Business Ethics* magazine over the period 2000-2007. Specifically, they study the stock price reaction to the press releases and the long-term return performance of the SRI portfolios. On the one hand, they find that new stocks that are included in the annual listing generate significant positive abnormal returns on the press release date. On the other hand, they document that the top 100 stocks outperform the S&P500 over longer holding periods, although the results are not statistically significant. Brammer et al. (2009), using the same list and analysing SRI portfolio performance over the period 2000-2004, find that over the year following the announcement, stocks of the '100 Best

[8]Examples of these ratings agencies include KLD Research & Analytics (e.g., Kempf and Osthoff, 2007; Galema et al., 2008; Derwall et al., 2011; Borgers et al., 2013), Sustainalytics (e.g., Auer, 2016; Auer and Schuhmacher, 2016), Asset 4 (e.g., Halbritter and Dorfleitner, 2015, Gonenc and Scholtens, 2017), EIRIS (e.g., Brammer et al., 2006), Vigeo (e.g., Van de Velde et al., 2005), and Innovest (e.g., Derwall et al., 2005).

Corporate Citizens' yield negative abnormal returns. Nevertheless, they suggest this can be a consequence of other stock features since when controlling for these firm characteristics the poor performance of the highly rated firms declines. Moreover, they find that companies in the top 100 but outside the S&P 500 can provide considerable positive abnormal returns.

Edmans (2011) analyses portfolios formed on the basis of the '100 Best Companies to Work For in America' in order to test the relationship between employee satisfaction and long-run stock portfolio returns. He shows that over the 1984-2009 period companies with stronger employee satisfaction not only have higher risk-adjusted returns but also exhibit both higher earnings announcement returns and higher long-term earnings surprises. The outperformance of high social rated stocks suggests that the stock market did not entirely value the intangible assets that companies created through strong relations with their employees. These results support those of Fulmer et al. (2003), who also investigate the link between employee relations and firms' performance using the '100 Best Companies to Work For in America'. Over the period 1995 to 2000, they find that the financial performance of the stocks in the list was better than that of a matched conventional peers sample. Similarly, Filbeck and Preece (2003) document that stocks in this list outperform a matched sample portfolio of conventional stocks over the period 1987 to 1999. More recently, Carvalho and Areal (2016) investigate the performance of portfolios of stocks listed on the '100 Best Companies to Work for in America' in times of financial crises and find that their financial performance and systematic risk remain unaffected in bear markets.

In turn, Anginer and Statman (2010) analyse the performance of portfolios composed by *Fortune* magazine's annual list of 'America's Most Admired Companies' by testing the relation between corporate reputation and subsequent returns. Over the period 1983 to 2007, they document that low-ranked stocks outperform high-ranked stocks, and that stocks of firms moving up the reputation scale lag stocks of firms moving down the scale. Filbeck et al. (2013) investigate whether the fact of being listed on different public surveys of exceptional firms (*Fortune*'s 'Most Admired Companies' and '100 Best Companies to Work For'; *Business Ethics* 'Best Corporate Citizens'; and *Working Mother*'s '100 Best Companies for Working Mothers') adds value to a portfolio and find that companies listed on the 'Most Admired Companies' and the 'Best Corporate Citizens' rankings are the most influential ones.

Outside the US market, Brzeszczyński and McIntosh (2014) investigate whether UK stocks listed on the Global-100 yield higher returns than the FTSE100 and FTSE4Good indices over the period 2000-2010. They find that the returns of the UK-SRI portfolios are higher than the returns of both the FTSE100 and the FTSE4GOOD indexes, although the differences are not statistically significant.

The studies mentioned above suffer from some limitations. First, all of them are country-specific studies (US or UK). Given the social and demographic country-specific factors (Bauer and Smeets, 2015), SRI financial performance to retail investors should be extended to more regions. Second, except for Brzeszczyński and McIntosh (2014) and Carvalho and Areal (2016), previous studies do not analyse the market state effect on financial performance, despite recent research on the performance of SRI equity funds and SRI fixed-income funds suggesting that portfolio performance is market state dependant. This issue is even more pertinent considering that it is recognised in the literature (e.g., Jagannathan and Wang, 1996) that evaluating portfolio performance without allowing for time-varying risk will lead to biased estimates of performance. We apply a conditional approach that allows for time-varying risk and performance by incorporating a dummy variable to distinguish between periods of crisis (bear markets) and non-crisis (bull markets). Third, Brzeszczyński and McIntosh (2014) do not analyse the statistical difference between the performance of those socially screened portfolios and conventional investment portfolios.

3.3 Data

In this study, stocks perceived as socially responsible are those that are included in the Global-100 list. This list was initiated in February 2005 and is released annually to report the 100 most sustainable businesses in the world. It is managed by Corporate Knights, who also provides indexing solutions and market-beating portfolios. Global-100 firms are considered to be socially responsible because they demonstrate, within their industries, a high capacity to integrate environmental, social and governance (ESG) criteria in their activities.

We identify and analyse stocks included in the Global-100 from January 2005 to December 2014. Monthly discrete returns of all stocks are computed based on the total return series (in US dollars) collected from Thomson Reuters database. To evaluate the long-term performance of SRI portfolios, we use the calendar-time portfolio approach (as in Carvalho and Areal, 2016). This approach involves creating an equally-weighted

portfolio of the stocks included in the Global-100 list in each year. Portfolios are rebalanced annually at the end of the month in which a new list is announced - each January, before the World Economic Forum in Davos. The list is published on www.global100.org and can be accessed easily and free of charge by any investor interested in engaging in SRI investment strategies. Thus, SRI criteria can easily be included in investment decisions without having to implement a complex social selection process (e.g., screening and engagement).

This paper analyses the performance of international SRI portfolios of stocks belonging to the list. From 2005 to 2014, 26 countries are represented in the sample. Table 1 shows the country stock allocation of the Global-100 during the full sample period. We can observe how the UK and the US are the most weighted countries in the sample - 19.40% and 16.72%, respectively. In this sense, it appears justified that previous research had focused on these markets. However, the percentage of countries such as Japan (12.54%), Canada (6.27%), and Australia (5.67%), among others, motivates the analysis of the SRI phenomenon to retail investors on other countries. Furthermore, it is worthwhile noting that the highest percentage (32.54%) of companies is from continental Europe firms. Although other countries are less represented, it is also interesting to analyse them, since, for instance, the list includes firms of emerging markets such as Brazil, India, South Korea or Taiwan, reflecting the fact that firms engaging in SRI practices are not restricted to developed markets.

Table 3-1. Country stock allocation

This table presents the country stock allocation of the Global-100 lists during the full sample period. (January 2005 to December 2014). Figures are represented in percentage (%) of the total number of stocks. The *Continental Europe Countries* encompass the percentage of European countries excluding UK.

Country	%	Country	%
Australia	5.67	Japan	12.54
Austria	0.90	Netherlands	1.79
Belgium	1.19	Norway	1.79
Brazil	2.09	Portugal	0.60
Canada	6.27	Singapore	1.79
Denmark	1.79	South Africa	0.60
Finland	2.69	South Korea	0.30
France	5.97	Spain	2.39
Germany	5.07	Sweden	4.18
Hong Kong	0.60	Switzerland	2.69
India	0.90	Taiwan	0.60
Ireland	0.30	United Kingdom	19.40
Italy	1.19	United States	16.72
Continental Europe Countries	*32.54*		

The financial performance of the SRI portfolio is evaluated relative to the S&P Global 100 Index. This index represents the financial performance of the 100 most important stocks in equity markets at a global level. Specifically, these are the firms with the highest capitalization in the S&P Global 1200, and are considered global businesses as they earn a large portion of their income doing business in different countries. This index was chosen for several reasons. Lydenberg and White (2015) point out that benchmarks should be defined by region, size and sector, and consequently, to make a suitable comparison, both the benchmark and the firm sample must have similar features. On that basis, the scope of firms making up the S&P Global 100 Index is global, just like the scope of the Global-100. Also, the number of firms in the S&P Global 100 Index is the same as the Global-100. Their fundamental difference is precisely what we are looking for: i.e., the appeal of following SRI criteria versus capitalization criteria can be evaluated using the S&P Global 100 Index. While the Global-100 firms are rated for specific SRI requirements, the S&P Global 100 Index firms are selected in terms of their capitalization.

Table 3-2. Descriptive statistics

The full sample period is from January 2005 to December 2014. Mean is the monthly arithmetic mean return, SD is the standard deviation. Mean diff (SD diff) is the mean return (standard deviation) of Global-100 portfolio (Global) minus S&P Global 100 Index (S&P) with *p*-values on *t*-tests (*F*-test) of equality of means (standard deviations). Risk/Reward ratio is the total return divided by standard deviation. Portfolios are rebalanced annually at the end of the month in which a new list is announced.

	Mean				SD				Reward/Risk	
	Global	S&P	Mean diff	*t-test*	Global	S&P	SD diff	*F-test*	Global	S&P
2005	0.0077	0.0024	0.0054	*0.5201*	0.0279	0.0223	0.0056	*1.5614*	0.2778	0.1060
2006	0.0222	0.0133	0.0089	*0.9592*	0.0241	0.0213	0.0027	*1.2723*	0.9230	0.6236
2007	0.0052	0.0067	-0.0015	*-0.1287*	0.0290	0.0278	0.0012	*1.0875*	0.1795	0.2416
2008	-0.0404	-0.0407	0.0003	*0.0125*	0.0697	0.0609	0.0087	*1.3064*	-0.5798	-0.6683
2009	0.0282	0.0167	0.0115	*0.3917*	0.0710	0.0727	-0.0017	*1.0486*	0.3973	0.2301
2010	0.0119	0.0022	0.0097	*0.3583*	0.0685	0.0641	0.0043	*1.1389*	0.1741	0.0343
2011	-0.0129	-0.0054	-0.0075	*-0.3386*	0.0557	0.0527	0.0030	*1.1162*	-0.2319	-0.1026
2012	0.0169	0.0078	0.0091	*0.4988*	0.0494	0.0399	0.0094	*1.5296*	0.3422	0.1945
2013	0.0125	0.0162	-0.0037	*-0.2796*	0.0361	0.0291	0.0069	*1.5345*	0.3470	0.5582
2014	-0.0013	0.0001	-0.0014	*-0.1196*	0.0296	0.0253	0.0043	*1.3734*	-0.0431	0.0031
Full period	0.0050	0.0019	0.0031	*0.4890*	0.0461	0.0416	0.0045	*1.2096*	0.1088	0.0463

Descriptive statistics on the average monthly returns, standard deviation and risk/reward ratio for the Global-100 portfolio and S&P Global 100 Index are presented in Table 2. Although the Global-100 portfolio yields higher returns than the S&P Global 100 Index in more years, as well as in the full sample period, these differences are not statistically significant. As to standard deviation, the Global-100 portfolio presents higher levels of risk than the S&P Global 100 Index in the majority of cases. However, the risk/reward

ratio shows that the relation between return and risk (standard deviation in this case) is somewhat better for the Global-100 portfolio than the S&P Global 100 Index.

Transaction costs are not considered in this study for several reasons: (1) the ability of retail investors to seek and negotiate the most favourable and advantageous investment alternatives will determine the final outcome of each investor; (2) transaction costs depend on aspects such as the amount of funds available for investing or the broker that retail investors use; (3) transaction costs affect the returns for retail investors investing in Global-100 stocks and in the S&P Global 100 Index; and (4) recent studies (e.g., Auer and Schuhmacher, 2016) consider transactions costs and find that this does not alter their main conclusions. Brzeszczyński and McIntosh (2014) point out that transaction costs would have to be disproportionately high to explain performance differences between SRI and conventional investments. Explanations for this can be found by taking a closer look at the changes of the social ratings over time, as changes do not occur very often (Auer and Schuhmacher, 2016), and because trading occurs only once a year and transaction costs are likely relatively trivial (Brammer et al., 2009).

3.4 Methods

We examine portfolio financial performance with stock market-based measures, in line with Scholtens (2008), Edmans (2011), Derwall et al. (2011) and Carvalho and Areal (2016) among others. Several researchers (Barber and Lyon, 1997; Fama, 1998; Loughran and Ritter, 2000) have shown that the magnitude and sometimes even the sign of the long-run abnormal returns are sensitive to alternative measurement methodologies. To determine the sensitivity of our results, we examine the financial performance using several approaches.

3.4.1 Sharpe ratio and significance tests

The Sharpe ratio (1966) - the ratio of excess return to standard deviation - is undoubtedly one of the most commonly used investments performance measure. Thus, as a general measure of financial performance and given the well-known interpretation of its results, retail investors may be interested in comparing the performance of alternative investment strategies according to this measure. From two investment portfolios i and j whose excess returns over the risk-free rate at time t are r_{ti} and r_{tj}

respectively, a total of *T* return pairs (r_{1i}, r_{1j}), …, (r_{Ti}, r_{Tj}) are observed. The difference between two Sharpe ratios is given by $\Delta = \mathrm{Sh}_i - \mathrm{Sh}_j = \mu_i/\sigma_i^2 - \mu_j/\sigma_j^2$, where μ and σ^2 are the sample mean and standard deviation respectively. As the value of the Sharpe ratio is actually an estimate from historical return data, statistical inference is applied in order to compare the two indicators. For this purpose, previous studies (e.g., DeMiguel and Nogales, 2009; Gasbarro et al., 2007) used the test of Jobson and Korkie (1981) and the correction proposed by Memmel (2003). However, this test is not valid if the returns distribution is non-normal or if the observations are correlated over time, which are phenomena quite common on financial returns time series data. Recently, Ledoit and Wolf (2008), hereafter LW, propose a studentized time series bootstrap approach that works asymptotically and has satisfactory properties in finite samples. Previous literature (e.g., Hall, 1992; Lahiri, 2003) shows the enhanced inference accuracy of the studentized bootstrap over standard inference based on asymptotic normality. LW propose to test $H_0: \Delta = \mathrm{Sh}_i - \mathrm{Sh}_j = 0$ by inverting a bootstrap confidence interval. A two-sided bootstrap confidence interval with nominal level 1-α for Δ (true difference between the Sharpe ratios) is constructed and if the interval does not include zero, then H_0 is rejected at a nominal level α. Specifically, LW propose to construct a symmetric studentized time series bootstrap confidence interval. To this end, the two-sided distribution function of the studentized statistic is approximated through the bootstrap by $\mathrm{F}\left(\left|\hat{\Delta} - \Delta\right|/s(\hat{\Delta})\right) \approx \mathrm{F}\left(\left|\hat{\Delta}^* - \Delta\right|/s(\hat{\Delta}^*)\right)$, where Δ is the true difference between the Sharpe ratios, $\hat{\Delta}$ is the estimated difference computed from the original data, $s(\hat{\Delta})$ is a standard error for $\hat{\Delta}$ (also calculated from the original data), $\hat{\Delta}^*$ is the estimated difference computed from bootstrap data, and $s(\hat{\Delta}^*)$ is a standard error for $\hat{\Delta}^*$ (also calculated from bootstrap data). Letting $z^*_{|\cdot|,\lambda}$ be a λ quantile of $\mathrm{F}\left(\left|\hat{\Delta}^* - \Delta\right|/s(\hat{\Delta}^*)\right)$, a bootstrap 1-α confident interval for Δ is given by $\hat{\Delta} \pm z^*_{|\cdot|,1-\alpha}\, s(\hat{\Delta})$. LW note that with heavy-tailed data or data of a time series nature, this quantile will typically be somewhat larger than its standard normal counterpart (used in the traditional tests) in small to moderate samples, resulting in more conservative inferences. To generate the bootstrap data, we use the circular block bootstrap of Politis and Romano (1992), resampling blocks of pairs from the observed pairs (r_{ti}, r_{tj}), *t*=1,…, *T*, with replacement. Applying the studentized circular block bootstrap requires a choice of the block size *b* and LW propose to use the calibration procedure of Loh (1987). LW suggest that *M* = 5000

bootstrap sequences is sufficient for reliable inference. The standard error $s(\hat{\Delta})$ is calculated through kernel estimation, specifically, the prewhitened quadratic spectral kernel of Andrews and Monahan (1992). The standard error $s(\hat{\Delta}^*)$ is the natural standard error calculated from the bootstrap data, making use of special block dependence structure. The bootstrap *p*-values are computed as $PV = \{\tilde{d}^{*,m} \geq d\} + 1/M + 1$, where $d = |\hat{\Delta}|/s(\hat{\Delta})$, the original studentized test statistic, $\tilde{d}^{*,m} = |\hat{\Delta}^{*,m} + \hat{\Delta}|/s(\hat{\Delta}^{*,m})$, denote the centered studentized statistic computed form the *m*th bootstrap sample by $d^{*,m}$, *m*=1,…, *M*, and *M* is the number of bootstrap resamples.

3.4.2 Firm features and systematic risk

Another approach to evaluate portfolio performance involves computing alphas from multi-factor models, as in Galema et al. (2008), Brammer et al. (2006), Edmans (2011), Derwall et al. (2011) and Humphrey et al. (2012). We examine performance using the four-factor Carhart (1997) model that captures the risk premiums associated with size and value/growth (as in Fama and French, 1993) as well as momentum, motivated by Jegadeesh and Titman (1993). The Carhart (1997) four-factor model is expressed by:

$$R_{p,t} - R_{f,t} = \alpha_p + \beta_{RMRF} RMRF_t + \beta_{SMB} SMB_t + \beta_{HML} HML_t + \beta_{MOM} MOM_t + \varepsilon_{p,t}$$

(Eq. 3-1)

where $R_{p,t}$ is the return of portfolio *p* on time *t*, $R_{f,t}$ is the risk-free rate and α_p is the estimated performance measure of the portfolio. In relation to the risk factors, $RMRF_t$ represents market excess returns (relative to the risk-free rate) on time *t*; SMB_t is the difference between the returns on diversified portfolios of small stocks and large stocks; HML_t is the difference between the returns on diversified portfolios of high book-to-market (value) stocks and low book-to-market (growth) stocks; and MOM_t is the difference between the returns on diversified portfolios of winning and losing stocks in the past year. The betas in the model represent the estimated risk measures associated to the risk factors: market, size, value/growth and momentum. Finally, $\varepsilon_{p,t}$ represents the residuals. To construct SMB and HML portfolios, we follow the recent Ferruz and Badía (2017) procedure, hereafter FB. The authors note that Fama and French (1993) construct portfolios once a year and maintain them invariable during the full year. However, variations in the characteristics of firms can occur during any given 12-month

period, which is not accounted by the Fama and French procedure. Taking month-to-month data and rebuilding the value and size portfolios at the end of each month, FB propose a more dynamic procedure that enhances the ability of the risk-factors and the model. To construct the MOM portfolio, we use six value-weighted portfolios formed on size and prior (2-12) returns. The portfolios are the intersections of two portfolios formed on size and three portfolios formed on prior (2-12) return. The MOM factor is also rebuilt at the end of each month. The monthly size breakpoint is the median market equity and the monthly prior return breakpoints are the 30^{th} and 70^{th} percentiles. Thus, MOM is the average return on the two high prior return portfolios (winners) minus the average return on the two low prior return portfolios (losers).

3.4.3 Geographical analysis

As outlined above, besides analysing performance at the global level, SRI financial performance is analysed at the regional level. Our international sample includes firms from 26 countries. Considering that a country-specific analysis would imply in some cases small samples, we form portfolios at the regional level. Following the MSCI market allocation, we analyse five regions (portfolios): (I) North America, that includes the United States and Canada; (II) Europe (except UK), that includes Austria, Belgium, Denmark, Finland, France, Germany, Ireland, Italy, the Netherlands, Norway, Portugal, Spain, Sweden and Switzerland; (III) United Kingdom; (IV) the Pacific region, that includes Australia, Hong Kong, Japan, New Zealand and Singapore; and (V) Emerging markets, that includes Brazil, India, South Africa, South Korea and Taiwan. This allocation is akin to the one of Fama and French (1998, 2012) who group countries in regions mainly by geographic location and market integration. Across the sample period, the average number of stocks in portfolio I (North America) is 22, in portfolio II (Europe ex-UK) 36, in portfolio III (UK) 20, in portfolio IV (Pacific regions) 16, and in portfolio V (Emerging markets) 7. We analyse UK firms and continental Europe firms separately not only due to the weight of the UK in the full sample, but also because of the differences of the UK market relative to continental Europe.[9] Separating UK firms from the Europe portfolio allows us to observe the SRI phenomenon on the UK market and to compare our results with previous studies.

[9]The UK financial market is recognized as being more similar to the US than to other continental European markets (Cernat, 2004).

3.4.4 Identification of different market states

Additionally, we analyse the financial performance of SRI portfolios in different market states. Recent research shows that different market states (such as expansion and recession periods), affect the performance of SRI equity funds and SRI fixed-income funds. We start by identifying the different market states across our sample period using the Pagan and Sossounov (2003), hereafter PS, approach.[10] PS develop a statistical approach to determine the peaks and troughs of a stock market index. A peak is identified at t time if the event $PK = [\ln P_{t-8}, \dots, \ln P_{t-1} < \ln P_t > \ln P_{t+1}, \dots, \ln P_{t+8}]$ occurs, where P_t represents the quotation of the relevant stock index, and a trough at time t if the event $TH = [\ln P_{t-8}, \dots, \ln P_{t-1} > \ln P_t < \ln P_{t+1}, \dots, \ln P_{t+8}]$ occurs. Consistent with the literature, we identify bear periods as those with a downtrend in the relevant stock market index of at least 20% from peak to trough. The MSCI ACWI[11] is used as the relevant stock market index since it is a coherent and complete representation of the market that captures the full spectrum of the global equity opportunity set without home bias. The index includes stocks across 23 developed markets and 23 emerging markets. With 2,480 constituents, the index covers approximately 85% of the global investable equity opportunity set. Table 3 shows the global bear market periods (Global-ACWI) identified over the period 2005-2014. The remaining periods are considered bull market periods. However, since this paper examines international socially responsible stock returns, we have to be cautious establishing unique global market states. Considering the different geographic areas of analysis, we thus proceed to identify different market states at the regional level. The relevant stock market indexes used are: the MSCI North America Index (portfolio I: North America); the MSCI Europe ex UK Index (portfolio II: Europe except UK); the MSCI United Kingdom Index (portfolio III: United Kingdom); the MSCI Pacific Index (portfolio IV: Pacific); and the MSCI Emerging Markets ex China Index[12] (portfolio V:

[10]This procedure to identify bull and bear markets is used for instance by Lee et al. (2013), and Carvalho and Areal (2016).
[11]Index prices are in USD. Data is obtained from www.msci.com.
[12]The MSCI Emerging Markets Index includes China as the most representative country. We use the MSCI Emerging Markets ex China Index since China is not included in our sample. Furthermore, the most representative countries in this index are those included in our sample: South Korea 20.62%, Taiwan 16.79%, India 12.11%, Brazil 10.43%, and South Africa 9.09%. Anyhow, we computed the analysis with both indices and obtained exactly the same results.

Emerging markets).[13] The regional bear periods are showed in Table 3. The remaining periods are considered as bull periods.

As expected, the downtrend associated to the international financial crisis (from 2007 to 2009) is observed both at the global and regional levels. Furthermore, we observe another bear market period in Europe ex-UK from May 2011 to May 2012, which can be associated to the Euro sovereign debt crisis, as well a bear market period in emerging markets (from May 2011 to September 2011), possibly due to financial contagion of fiscal risks in the US and sovereign debt sustainability in Europe.

Table 3-3. Bear market states

This table identifies periods of bear market according to the Pagan and Sossounov (2003) procedure. The sample period studied is from January 2005 to December 2014. The indices used are the MSCI ACWI Index (Global); the MSCI North America Index (portfolio I: North America); the MSCI Europe ex UK Index (portfolio II: Europe except UK); the MSCI United Kingdom Index (portfolio III: United Kingdom); the MSCI Pacific Index (portfolio IV: Pacific); and the MSCI Emerging Markets ex China Index (portfolio V: Emerging markets). Consistent with literature, we require the rise (fall) of the market being greater (less) than either 20%. We test the window breadth for eight, nine and ten months and obtain the same results.

Portfolio	Start date	Index value (Points)	End date	Index value (Points)	Change in market index	Length of bear period (months)
Global-ACWI	2007/11	408.105	2009/02	187.168	-0.5414	16
North America	2007/11	1 558.805	2009/02	776.949	-0.5016	16
Europe except UK	2007/11	2 452.294	2009/02	985.823	-0.5980	16
	2011/05	1 794.745	2012/05	1 231.996	-0.3472	13
UK	2007/11	1 638.644	2009/02	672.550	-0.5896	16
Pacific	2007/11	2 763.476	2009/02	1 369.571	-0.5044	16
Emerging Markets	2007/11	4 030.146	2009/02	1 610.415	-0.6004	16
	2011/05	3 945.570	2011/09	3 011.914	-0.2366	5

3.4.5 Performance in different market states

To analyse the market state effect on financial performance we use a conditional four-factor model with dummy variables, in the spirit of Nofsinger and Varma (2014) and Leite and Cortez (2015). Our model allows risk and performance to vary across different market states by incorporating two dummy variables, as follows:

[13]Prices for all indices are in USD. Data is obtained from www.msci.com. Indices used for the remaining regions cover the same countries as our regional portfolios. The MSCI North America Index covers US and Canadian firms; the MSCI Europe ex UK Index covers firms from Austria, Belgium, Denmark, Finland, France, Germany, Ireland, Italy, the Netherlands, Norway, Portugal, Spain, Sweden and Switzerland; the MSCI United Kingdom Index covers stocks from UK; and the MSCI Pacific Index covers firms from Australia, Hong Kong, Japan, New Zealand and Singapore.

$$R_{p,t} - R_{f,t} = \alpha_{Bear} D_{Bear,t} + \alpha_{Bull} D_{Bull,t} + \beta_{1Bear} RMRF_t D_{Bear,t} + \beta_{1Bull} RMRF_t D_{Bull,t} + \beta_{2Bear} SMB_t D_{Bear,t} + \beta_{2Bull} SMB_t D_{Bull,t} + \beta_{3Bear} HML_t D_{Bear,t} + \beta_{3Bull} HML_t D_{Bull,t} + \beta_{4Bear} MOM_t D_{Bear,t} + \beta_{4Bull} MOM_t D_{Bull,t} + \varepsilon_{p,t}$$

(Eq. 3-2)

Where $D_{Bear,t}$ is a dummy variable that takes value of one for bear market periods and zero otherwise and $D_{Bull,t}$ is a dummy variable that takes value of one for bull market periods and zero otherwise; α_{Bear} corresponds to the financial performance in bear markets and α_{Bull} in bull markets; β_{1Bear}, β_{2Bear}, β_{3Bear} and β_{4Bear} correspond to the factor loadings in bear periods; and β_{1Bull}, β_{2Bull}, β_{3Bull} and β_{4Bull} in bull periods. It is important to note this our model differs from that of Nofsinger and Varma (2014) by incorporating the dummy variables not only in the alphas but also in the risk factors, thereby enabling the analysis of financial performance and risk exposures in different market states.

3.5 Empirical results

3.5.1 SRI and conventional portfolio performance

This section presents the empirical results. Table 4 shows the results of applying the Sharpe ratio and the LW procedure to estimate the statistical significance of the difference between the Sharpe ratio of the SRI portfolio (Global-100 stocks) and conventional investments (S&P Global 100 Index), as well as the results of estimating the four-factor Carhart (1997) model to both portfolios. Furthermore, in order to investigate the differences in financial performance between both portfolios, we also estimate the alphas of a portfolio constructed by subtracting the returns of the S&P Global 100 Index from the returns of the Global-100 portfolio ('difference' portfolio).

Considering the full sample period, the Sharpe estimate for the Global-100 portfolio is 0.0751 and for the S&P Global 100 Index 0.0162, resulting in a difference of 0.0589. The LW test produces a *p*-value of 0.0569, meaning that the difference between the Sharpe ratio of both portfolios is statistically significant. These results are supported by the alpha estimates. The Global-100 portfolio shows a positive and significant alpha and the S&P Global 100 Index yields a negative although not statistically significant alpha.

Table 3-4. Portfolio financial performance and risk estimates

This table shows estimates of performance and risk for the Global 100 portfolio (Global) and the S&P Global 100 Index (S&P). Diff is the portfolio constructed by subtracting the returns of the S&P Global 100 Index from the returns of the Global-100 portfolio. The full sample period is from January 2005 to December 2014. Portfolio performance is evaluated by means of the Sharpe ratio and the alpha from the four-factor Carhart (1997) model. The LW procedure is used to identify statistical significant differences between the Sharpe ratio of both portfolios, and values in brackets represent the *p*-value for equal Sharpe ratios. The Carhart (1997) model is estimated by OLS based on the heteroskedasticity and autocorrelation adjusted errors of Newey and West (1987). The SMB and HML factors are constructed following FB and MOM following the Carhart (1997) approach. The MSCI ACWI Index is the market benchmark in the Carhart (1997) model. One-month US T-bills proxy for the risk-free rate. R2 Adj. is the adjusted coefficient of determination. Values in parenthesis are the *t*-statistics. The asterisks are used to represent the statistically significant coefficients at the 1% (***), 5% (**) and 10% (*) significance levels.

	Sharpe	Alpha	RMRF	SMB	HML	MOM	R2 Adj.
Global	0.0751	0.0025**	0.9843***	0.1315***	0.2351***	-0.0194	0.9614
		(2.4432)	(37.0308)	(4.1246)	(4.8310)	(-0.7825)	
S&P	0.0162	-0.0010	0.9432***	0.1153***	-0.1250**	0.0169	0.9530
		(-0.9180)	(31.6051)	(4.2158)	(-2.1312)	(0.5718)	
Diff	0.0589*	0.0034**	0.0411	0.0162	0.3600***	-0.0363	0.4408
	[0.0569]	(2.2868)	(0.9930)	(0.3771)	(4.8316)	(-0.8797)	

The difference in performance between both portfolios, measured by the alpha of the 'difference' portfolio, is statistically significant, indicating that the Global 100 portfolio outperforms the S&P Global 100 Index. Thus, both financial performance measures indicate statistically significant differences between SRI and conventional investments, suggesting that the Global 100 portfolio yields better financial performance than the S&P Global 100 Index. As to risk factors, both portfolios show a positive and statistically significant exposure to the size factor, reflecting a tendency for the portfolios to be exposed to smaller firms. Furthermore, the Global 100 portfolio presents a significant positive loading on the value factor, whereas the S&P Global 100 Index has a significant and negative exposure to this factor. Considering the results of the 'difference' portfolio, we can conclude that the SRI portfolio is significantly more exposed to value stocks. Regarding the momentum factor, we do not find any statistically significant coefficients. Overall, our results are in line with previous studies such as Filbeck (2009), Edmans (2011) and Filbeck (2013), and suggest that socially responsible investors are able to benefit from the outperformance of a SRI strategy relative to conventional investments.

The results on the portfolio performance of the SRI portfolios at the regional level are presented in Table 5. Estimates of the Sharpe ratio and four-factor model for each region are reported. With respect to the Sharpe ratios, three portfolios show positive values and two other show negative values for this measure. If the portfolios are ranked by the Sharpe values, portfolio P1 (North America) yields the highest financial

performance, followed by portfolio P2 (Europe ex-UK). Portfolio P5 (Emerging markets) obtains the lowest financial performance, followed by portfolio P4 (Pacific region). The alpha estimates further allows us to explore the portfolio performance, controlling for the four risk factors. Portfolios P1 and P2 yield a positive and statistically significant alpha (at the 1% level); portfolios P3 and P4 show insignificant alphas, and portfolio P5 shows a marginal (at the 10% level) statistically significant negative alpha. These results suggest that the significant differences observed at the global level between the Global-100 portfolio and the S&P Global 100 Index are driven mainly by portfolios P1 and P2. On the other hand, it is possible to observe how risk sensitivities oscillate notably among regions. The size factor loses relevance in Pacific and Emerging markets; the value factor is only significant in the North America and UK regions; and the momentum effect is documented solely in the North American portfolio. Thus, the typical risk factors seem to present a limited capacity to explain some specific regional portfolio returns. Brzeszczyński and McIntosh (2014) document that the returns of the UK-SRI portfolios cannot be consistently explained by conventional factors other than the market factor. However, in contrast, our size and value risk factors constructed via FB are significant for this region. When analysing North American socially responsible stocks, Brammer et al. (2006) find negative loadings on the market, size, value, and momentum factors, although only size and momentum are statistically significant. In contrast, our results for portfolio P1 (North America) point out significant positive loadings on the market, size and value factors, and a significant negative exposure on momentum. Constructing the size, value and momentum portfolios following the FB procedure seems to have a positive influence on the significance of the risk factors. As to the financial performance, our results are in line with previous evidence for the UK market (e.g., Humphrey et al., 2012; Brzeszczyński and McIntosh, 2014) and for the US market (e.g., Edmans, 2011; Filbeck, 2013), and are in contrast with Brammer et al. (2006) and Mollet and Ziegles (2014). Since this study is, as far as we are aware of, the first to analyse SRI portfolios focused on retail investor possibilities in pacific and emerging markets, our results are novel for this geography.

Table 3-5. SRI Financial performance and risk at the regional level

This table shows estimates of performance and risk for each regional portfolio. Five regional portfolios are constructed: P1 corresponds to North America; P2 is Europe ex-UK; P3 is UK; P4 is Pacific; and P5 correspond to Emerging markets. The full sample period is from January 2005 to December 2014. The estimates for the P5 portfolio start in January 2010, considering previously there are no stocks from this region in the sample. Portfolio performance is evaluated by means of the Sharpe ratio and the alpha from the four-factor Carhart (1997) model. The Carhart (1997) model is estimated by OLS based on the heteroskedasticity and autocorrelation adjusted errors of Newey and West (1987). Portfolios SMB and HML are constructed for each region specifically following FB and MOM following the Carhart approach. Market benchmarks are the MSCI North America for P1; the MSCI Europe ex UK for P2; the MSCI United Kingdom for P3; the MSCI Pacific for P4, and Emerging markets ex china for P5. One-month US T-bills proxy for the risk-free rate. R2 Adj. is the adjusted coefficient of determination. Values in parenthesis are the *t*-statistics. The asterisks are used to represent the statistically significant coefficients at the 1% (***), 5% (**) and 10% (*) significance levels.

	Sharpe	Alpha	RMRF	SMB	HML	MOM	R2. Adj.
P1	0.1560	0.0033***	0.9247***	0.2486***	0.0742***	-0.0662***	0.9710
		(2.7030)	(37.4582)	(6.4825)	(3.1980)	(-3.5201)	
P2	0.0738	0.0033***	0.9566***	0.2297***	0.0528	-0.0353	0.9700
		(3.2330)	(42.9523)	(4.7386)	(1.5436)	(-1.0756)	
P3	0.0132	0.0024	0.8650***	0.2585***	0.1164**	-0.0931	0.8988
		(1.4895)	(19.5529)	(4.2252)	(2.1344)	(-1.5885)	
P4	-0.0275	-0.0025	1.0646***	0.0043	0.0420	-0.0280	0.8967
		(-1.5309)	(34.0875)	(0.0874)	(1.2782)	(-1.0155)	
P5	-0.0713	-0.0182*	1.1944***	0.0238	0.1256	-0.0320	0.8722
		(-1.8718)	(24.7482)	(0.3319)	(1.4017)	(-0.4880)	

The statistical differences between Sharpe ratios by pairs of regions, as well as the alpha of the 'differences' portfolio, also between pairs of regions, are presented in table 6. In the up-right side of the table, we can observe that the differences between the Sharpe ratios of portfolio P1 are statistically significant from those of portfolios P4 and P5. Portfolio P2 also shows statistically significant differences in relation to the Sharpe ratios of portfolios P4 and P5, whereas portfolio P3 yields a Sharpe ratio that is significantly different from that of portfolio P5. The difference between portfolios P4 and P5 is not significant. Furthermore, in the down-left side of the table, we present the alpha estimates of the difference portfolios between pairs of regions. We can observe that, after controlling for four risk-factors, the alpha of portfolio P1 is statistically different from all other portfolios; the alpha of portfolio P2 is statistically different in relation to portfolios P3 and P5; and the alphas of portfolios P3 and P4 are statistically significant different from that of portfolio P5. These results complement the results presented in table 5. It appears that portfolios P1 and P2 are the main drivers of SRI financial performance. Given the statistical differences in financial performances among regional portfolios, in line with previous studies (e.g., Nilsson, 2008; Heimann et al., 2011; Hörisch et al., 2015; Bauer and Smeets, 2015), these results suggest country-

specific factors may affect the relationship between corporate social and financial performance.

Table 3-6. Differences in SRI financial performance and risk at the regional level

This table shows financial performance differences between regional portfolios. Up-right side of the table presents the Sharpe differences between pairs of regions. The LW procedure is used to identify statistical significant differences between the Sharpe ratio of pairs of regional portfolios. Down-left side of the table shows the alpha estimates of the difference portfolios between pairs of regions. Difference portfolios are constructed by subtracting the returns of a regional portfolio from the returns of another one. Alphas are estimated by the four-factor Carhart (1997) model. This model is regressed by OLS based on the heteroskedasticity and autocorrelation adjusted errors of Newey and West (1987). P1 corresponds to the North America portfolio; P2 to Europe ex-UK; P3 to the UK; P4 to Pacific; and P5 to Emerging markets. The full sample period is from January 2005 to December 2014. Differences with the P5 portfolio are estimated from January 2010, considering previously there are no stocks from this region in the sample. The asterisks are used to represent the statistically significant coefficients at the 1% (***), 5% (**) and 10% (*) significance levels.

	P1	P2	P3	P4	P5
P1	--	0.0822	0.1428	0.1836**	0.2773***
P2	0.0148***	--	0.0605	0.1013**	0.1682*
P3	0.0216***	0.0068*	--	0.0408	0.2067**
P4	0.0191***	0.0044	-0.0024	--	0.1146
P5	0.0140***	0.0129*	0.0135**	0.0191***	--

Finally, table 7 presents estimates of performance and risk of the Global-100 portfolio, the S&P Global 100 Index, as well as the regional portfolios, across different market states. In panel A, we observe that in bear markets the alpha is negative, although not statistically significant, for both portfolios, indicating a neutral performance. During bull market periods, the Global-100 portfolio yields a positive and statistically significant alpha whereas the S&P Global 100 index shows a negative and marginal (at the 10% level) statistically significant alpha. In bull markets, the alpha of the 'difference' portfolio is statistically significant, showing an outperformance of the Global-100 portfolio relative to the S&P Global 100 index. In bear market periods, there are no statistical significant differences between the performance of both portfolios. Brzeszczyński and McIntosh (2014) show that SRI stock portfolios yield higher mean returns than conventional benchmarks during bull and bear market periods in the UK market, although the differences are small and not statistically significant. By means of a more robust methodology, Carvalho and Areal (2016) find that the financial performance of socially responsible stocks is not affected during bear market periods. We document that the SRI portfolio shows neutral performance in bear market periods and a positive performance in bull markets. Furthermore, we also show that the SRI portfolio outperforms the conventional portfolio in bull market periods.

Table 3-7. Financial performance in different market states

This table presents estimates of performance and risk of the Global-100 portfolio, the S&P Global 100 Index, as well as the regional portfolios, in different market states, based on the conditional model (equation 2). The model is estimated by OLS based on the heteroskedasticity and autocorrelation adjusted errors of Newey and West (1987). The Pagan and Sossounov (2003) procedure is used in order to identify the different market states (bear and bull). G (S&P) corresponds to the Global-100 portfolio (S&P Global 100 index); P1 corresponds to the North America portfolio; P2 to Europe ex-UK; P3 to the UK; P4 to Pacific; and P5 to Emerging markets. Diff is the portfolio constructed by subtracting the returns of the S&P Global 100 Index from the returns of the Global-100 portfolio. The coefficients β_1, β_2, β_3 and β_4 represent the factor loadings on the market excess return, size, value and momentum factors, respectively. The full sample period is from January 2005 to December 2014. The estimates for the P5 portfolio start in January 2010, considering previously there are no stocks from this region in the sample, therefore, only the second bear market period is studied. R2 Adj. is the adjusted coefficient of determination. Values in parenthesis are the *t*-statistics. The asterisks are used to represent the statistically significant coefficients at the 1% (***), 5% (**) and 10% (*) significance levels.

Panel A: The Global-100 portfolio and the S&P Global 100 Index.											
	α_{Bear}	α_{Bull}	β_{1Bear}	β_{1Bull}	β_{2Bear}	β_{2Bull}	β_{3Bear}	β_{3Bull}	β_{4Bear}	β_{4Bull}	R2 Adj.
G	-0.0049	0.0028**	0.8822***	1.0241***	0.3536***	0.2583***	-0.0683	0.1753***	-0.1506***	0.0183	0.9667
	(-1.1553)	(2.4929)	(12.5231)	(37.8883)	(3.5519)	(4.6964)	(-1.3747)	(4.5157)	(-3.8889)	(0.5639)	
S&P	-0.0040	-0.0024*	0.8816***	0.9926***	-0.1916	-0.2117***	0.2111***	0.0471	0.0686	0.0166	0.9551
	(-1.3350)	(-1.9430)	(18.4066)	(31.9604)	(-1.3656)	(-3.4536)	(3.8732)	(1.1775)	(1.3341)	(0.5940)	
Diff	-0.0010	0.0051***	0.0006	0.0315	0.5453***	0.4700***	-0.2794***	0.1283**	-0.2191***	0.0017	0.4837
	(-0.1505)	(3.3375)	(0.0051)	(0.9319)	(3.4629)	(6.3578)	(-4.0557)	(2.0921)	(-3.1430)	(0.0379)	
Panel B: Regional SRI portfolios.											
	α_{Bear}	α_{Bull}	β_{1Bear}	β_{1Bull}	β_{2Bear}	β_{2Bull}	β_{3Bear}	β_{3Bull}	β_{4Bear}	β_{4Bull}	R2 Adj.
P1	-0.0025	0.0030**	0.7918***	0.9693***	0.1832**	0.3112***	0.0572*	0.0610**	-0.1268***	-0.0632***	0.9726
	(-0.4074)	(2.3875)	(7.3994)	(33.1152)	(2.4863)	(6.3660)	(1.8707)	(2.1757)	(-4.2221)	(-3.2156)	
P2	0.0004	0.0025*	0.9765***	0.9827***	0.0804	0.3224***	-0.0354	0.0761*	-0.0993***	0.0043	0.9723
	(0.2306)	(1.9502)	(22.7625)	(33.8107)	(0.7803)	(7.4091)	(-1.0696)	(1.8160)	(-3.0631)	(0.1191)	
P3	-0.0100	0.0043***	0.7833***	0.8236***	0.5750***	0.2285***	0.0265	0.1407***	-0.1535**	0.0458	0.9188
	(-1.3206)	(2.8773)	(10.5601)	(18.7884)	(4.4992)	(4.4297)	(0.4112)	(2.8883)	(-2.5930)	(1.1700)	
P4	-0.0119*	-0.0020	1.0495***	1.0701***	0.1557**	0.0378	0.3365***	0.0172	-0.2432***	-0.0118	0.9034
	(-1.7764)	(-1.1310)	(28.3980)	(22.8302)	(2.5524)	(1.0395)	(2.8888)	(0.5827)	(-4.1649)	(-0.3685)	
P5	0.0059***	-0.0129	-0.0280***	1.2431***	0.1911***	0.0192	-0.5914***	0.0650	-0.3284***	-0.0348	0.8703
	(0.0000)	(-1.1872)	(0.0000)	(23.2819)	(0.0000)	(0.2339)	(0.0000)	(0.6633)	(0.0000)	(-0.4922)	

The outperformance of the Global-100 portfolio in relation to the S&P Global 100 Index during bull markets seems to be related to the higher exposure to the size and value factors. Panel B shows the performance and risk of the regional portfolios in different market states. Portfolios P1, P2 and P3 exhibit positive and statistically significant alphas in up markets and present a similar exposure to risk factors during these periods. All portfolios have a significant positive exposure to the size and value factors, although momentum is only significant and negative for portfolio P1. The negative exposure to momentum may be related to the more narrowed investment universe of SRI (Leite and Cortez, 2015). In contrast, portfolio P4 shows a marginal negative and statistically significant alpha in bear markets. This portfolio tends to be positively exposed to the size and value factors, and negatively exposed to momentum. Portfolio P5 is the only one showing a positive and statistically significant alpha in down markets periods, related to a significant positive exposure to the size factor and a

significant negative exposure to the value and momentum factors. The regional analysis of performance in different market states allows us to conclude that the performance of the Global-100 portfolio is mostly influenced by regional portfolios P1, P2 and P3. Overall, the results reinforce the argument in favour of country-specific features on the relationship between corporate social and financial performance.

3.5.2 Robustness checks

Finally, we report a variety of supplementary checks in order to verify the robustness of our results. First, alternative risk-free rates are used for the calculation of excess returns. Specifically, we calculate the excess returns using the 1-month European Interbank Offered Rate (EURIBOR) and the UK 1-month T-bill as the risk-free rates. Statistical significant financial performance differences between the Global-100 portfolio and the S&P Global 100 Index are even higher (p-value < 0.01) using alternative risk-free rates[14]. Second, other commonly used global indices are employed as alternative conventional investment benchmarks. We consider the Russell Global Index, the Thomson Reuters Global Index, the S&P Global 1200 Index, the STOXX Global 1800 Index, the World DataStream Market Index, and the FTSE Global Index, and assess the financial performance differences between the Global-100 portfolio and these alternative benchmarks. Statistically significant differences between the 1% and the 10% level are found for all 'differences' portfolios, after controlling for the risk factors. Hence, again, there is strong evidence on the outperformance of SRI portfolios relative to conventional investments. Finally, other financial performance evaluation measures are considered. We employ the modification proposed by Ferruz and Sarto (2004) regarding the Sharpe ratio (1966) used in studies such as Scholz (2007) and Luo et al. (2015). Ferruz and Sarto (2004) note that the Sharpe ratio assumes positive portfolio excess returns. However, this is not always the case. Consequently, when this happens, the Sharpe ratio can present anomalous results. In this context, Ferruz and Sarto (2004) propose a correction to the Sharpe ratio, as follows: $FS_{p,t} = \left(R_{p,t}/R_{f,t}\right)/\sigma_{p,t}$, where $R_{p,t}$ is the portfolio p return on time t, $R_{f,t}$ is the risk-free return on time t, and $\sigma_{p,t}$ is standard deviation of the portfolio p on time t. We also employ the Sortino ratio (Sortino and van der Meer, 1991, Sortino and Price, 1994), used by authors such as Leggio and Lienv (2003), Meligkotsidou et al. (2009) and Auer (2016) to measure

[14]The specific results of this section are not presented for the sake of brevity and because our main results and conclusions are not altered. Nonetheless, detailed results are available upon request.

performance on the basis of the lower partial moments (LPM). According to the Sortino ratio, risk is measured by the negative deviations of returns in relation to a minimum acceptable return (e.g., zero, the risk-free rate or the average return). In our case, we use a rolling interest rate based on the evolution of the risk-free monthly interest rate. The Sortino specification is $S_{p,t} = R_{p,t} - \varphi / \left(\frac{1}{T}\sum_{t=1}^{T} \max[\varphi - R_{p,t}, 0]^2\right)^{1/2}$, where $R_{p,t}$ is the portfolio *p* return on time *t*, and φ is the target return or minimum acceptable return. Using these performance measures, we find consistent results. As to the Ferruz and Sarto (2004) correction, the Global-100 portfolio shows a value greater than twice of that of the S&P Global 100 Index. When we analyse the performance using the LPM, the difference is even higher.

3.6 Conclusions

In recent periods there has been a considerable increase in the popularity of SRI among retail investors. Moreover, the technological developments in trading systems, reducing transaction costs and commissions, have encouraged retail investors' trading. The impact of considering social criteria on the performance of SRI portfolios is therefore an important issue for retail investors. Previous evidence on the relation between SRI and portfolio financial performance is extensive. Yet, most studies are conducted from the perspective of institutional investors and not from the perspective of retail investors who wish to construct SRI portfolios. Research on the performance of SRI portfolios constructed on the basis of free and available information to investors, which may be useful to retail investors, is somewhat scarce, and focuses mainly the US and the UK markets.

This paper highlights this issue and analyses the performance of SRI portfolios constructed on the basis of the Global 100 list over the period 2005 to 2014. Since previous evidence is focused on specific countries, we provide evidence of SRI financial performance at the worldwide level as well as at the regional level, for 5 regions (North America, Europe except UK, United Kingdom, Pacific region and Emerging markets). Additionally, since recent research shows that SRI performance can differ across market states, we analyse SRI portfolio performance in periods of bull and bear markets.

Our results show that the Global-100 portfolio outperforms the S&P Global 100 Index. In terms of investment styles, both SRI and conventional investments are more exposed

to small firms, whereas SRI is more associated to value firms and conventional investments to a growth stocks. The results on SRI financial performance and risk at the regional level show statistical differences in the financial performance of the SRI regional portfolios. The regional analysis allows us to conclude that the performance of the Global-100 portfolio is mostly influenced by three specific regional portfolios: North America and Europe ex-UK (positive impact) and emerging markets (negative impact). Thus, our results suggest that country-specific factors may affect the relationship between corporate social and financial performance. Nevertheless, as a limitation of our study, we do not study the influence of concrete social factors in investment decisions. Risk sensitivities oscillate notably among regions and we find that the typical risk factors present a limited capacity to explain some specific regional portfolio returns. The analysis on the differences by pairs of regions highlights statistically significant differences among regional portfolios and further motivates the debate on the effect of country-specific factors in responsible investing. As to the differences in performance between SRI and conventional investments across different market states, the results show that the financial performance in bear market periods is neutral for both portfolios. In bull market periods, the Global-100 portfolio shows a positive and statistically significant performance whereas the S&P Global 100 index yields negative and marginal statistically significant financial performance. The Global-100 thus outperforms the S&P 100 Index in up markets. Furthermore, we document that this outperformance is related to a positive and statistically significant exposure to the size and value risk factors. The regional analysis in this context shows how the regions present miscellaneous exposures in different market states. Our results are robust to several test related to the use of alternatives risk-free rates, benchmarks indexes, and financial performance measures.

In sum, our empirical evidence indicates that socially conscious retail investors are able to implement a SRI strategy that outperforms the S&P Global 100. In addition, the different results uncovered at the regional level suggest that country-specific factors may affect the relationship between corporate social and financial performance. Finally, we document that social screened investments are not negatively affected in bad times, and that in good times their performance increases, outperforming comparable conventional investments. This study has been performed from a retail investor perspective, but, of course the results are also useful for institutional investors when constructing their SRI strategies.

References

Andrews, D. W., & Monahan, J. C. (1992). An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator. *Econometrica*, 953-966.

Anginer, D., & Statman, M. (2010). Stocks of admired and spurned companies. *The Journal of Portfolio Management, 36*(3), 71-77.

Auer, B. R. (2016). Do socially responsible investment policies add or destroy European stock portfolio value? *Journal of Business Ethics, 135*(2), 381.

Auer, B. R., & Schuhmacher, F. (2016). Do socially (ir) responsible investments pay? new evidence from international ESG data. *The Quarterly Review of Economics and Finance, 59*, 51-62.

Barber, B. M., & Lyon, J. D. (1997). Firm size, book-to-market ratio, and security returns: A holdout sample of financial firms. *The Journal of Finance, 52*(2), 875-883.

Bauer, R., & Smeets, P. (2015). Social identification and investment decisions. *Journal of Economic Behavior & Organization, 117*, 121-134.

Bauer, R., Koedijk, K., & Otten, R. (2005). International evidence on ethical mutual fund performance and investment style. *Journal of Banking & Finance, 29*(7), 1751-1767.

Becchetti, L., Ciciretti, R., Dalò, A., & Herzel, S. (2015). Socially responsible and conventional investment funds: Performance comparison and the global financial crisis. *Applied Economics, 47*(25), 2541-2562.

Benijts, T. (2010). A framework for comparing socially responsible investment markets: An analysis of the Dutch and Belgian retail markets. *Business Ethics: A European Review, 19*(1), 50-63.

Benson, K. L., & Humphrey, J. E. (2008). Socially responsible investment funds: Investor reaction to current and past returns. *Journal of Banking and Finance, 32*(9), 1850-1859.

Borgers, A., Derwall, J., Koedijk, K., & Horst, J. (2013). Stakeholder Relations and Stock Returns: On Errors in Investors' Expectations and Learnings. *Journal of Empirical Finance* , 22, 159-175.

Brammer, S., Brooks, C., & Pavelin, S. (2006). Corporate social performance and stock returns: UK evidence from disaggregate measures. *Financial Management, 35*(3), 97-116.

Brammer, S., Brooks, C., & Pavelin, S. (2009). The stock performance of America's 100 best corporate citizens. *The Quarterly Review of Economics and Finance, 49*(3), 1065-1080.

Brzeszczyski, J., & McIntosh, G. (2014). Performance of portfolios composed of British SRI stocks. *Journal of Business Ethics, 120*(3), 335.

Butt, H. A., & Virk, N. S. (2017). Momentum profits and time varying illiquidity effect. *Finance Research Letters, 20*, 253-259.

Capelle-Blancard, G., & Monjon, S. (2012). Trends in the literature on socially responsible investment: Looking for the keys under the lamppost. *Business Ethics: A European Review, 21*(3), 239-250.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance, 52*(1), 57-82.

Carvalho, A., & Areal, N. (2016). Great places to work®: Resilience in times of crisis. *Human Resource Management, 55*(3), 479-498.

Cernat, L. (2004). The emerging European corporate governance model: Anglo-Saxon, Continental, or still the century of diversity? *Journal of European Public Policy, 11*(1), 147-166.

Cortez, M. C., Silva, F., & Areal, N. (2009). The performance of European socially responsible funds. *Journal of Business Ethics, 87*(4), 573-588.
Cortez, M. C., Silva, F., & Areal, N. (2012). Socially responsible investing in the global market: The performance of US and European funds. *International Journal of Finance & Economics, 17*(3), 254-271.
DeMiguel, V., & Nogales, F. J. (2009). Portfolio selection with robust estimation. *Operations Research, 57*(3), 560-577.
Derwall, J., Guenster, N., Bauer, R., & Koedijk, K. (2005). The Eco-Efficiency Premium Puzzle. *Financial Analysts Journal, 61*(2), 51-63.
Derwall, J., Koedijk, K., & Ter Horst, J. (2011). A tale of values-driven and profit-seeking social investors. *Journal of Banking & Finance, 35*(8), 2137-2147.
Edmans, A. (2011). Does the stock market fully value intangibles? employee satisfaction and equity prices. *Journal of Financial Economics, 101*(3), 621-640.
Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics, 33*(1), 3-56.
Fama, E. F., & French, K. R. (1998). Value versus growth: The international evidence. *The Journal of Finance, 53*(6), 1975-1999.
Fama, E. F., & French, K. R. (2012). Size, value, and momentum in international stock returns. *Journal of Financial Economics, 105*(3), 457-472.
Ferruz, L., & Badía, G. (2017). Adapting and testing the Fama and French model, with some variations of company characteristics. *Applied Economics Letters, 24*(5), 342-345.
Ferruz, L., & Sarto, J. L. (2004). An analysis of Spanish investment fund performance: Some considerations concerning Sharpe's ratio. *Omega, 32*(4), 273-284.
Ferruz, L., Muñoz, F., & Vargas, M. (2012). Managerial abilities: Evidence from religious mutual fund managers. *Journal of Business Ethics, 105*(4), 503-517.
Filbeck, G., & Preece, D. (2003). Fortune's best 100 companies to work for in America: Do they work for shareholders? *Journal of Business Finance & Accounting, 30*(5-6), 771-797.
Filbeck, G., Gorman, R., & Zhao, X. (2009). The "Best corporate citizens": Are they good for their shareholders? *Financial Review, 44*(2), 239-262.
Filbeck, G., Gorman, R., & Zhao, X. (2013). Are the best of the best better than the rest? the effect of multiple rankings on company value. *Review of Quantitative Finance and Accounting, 41*(4), 695.
Fulmer, I. S., Gerhart, B., & Scott, K. S. (2003). Are the 100 best better? an empirical investigation of the relationship between being a "great place to work" and firm performance. *Personnel Psychology, 56*(4), 965-993.
Galema, R., Plantinga, A., & Scholtens, B. (2008). The stocks at stake: Return and risk in socially responsible investment. *Journal of Banking & Finance, 32*(12), 2646-2654.
Gasbarro, D., Wong, W., & Kenton Zumwalt, J. (2007). Stochastic dominance analysis of iShares. *The European Journal of Finance, 13*(1), 89-101.
Global Sustainable Investment Alliance – GSIA. (2016). *Global sustainable investment review 2016* .Accessed in July 2017, available at www.gsi-alliance.org
Gonenc, H., & Scholtens, B. (2017). Environmental and financial performance of fossil fuel firms: A closer inspection of their interaction. *Ecological Economics, 132*, 307-328.
Graafland, J. J., & van de Ven, Bert W. (2011). The credit crisis and the moral responsibility of professionals in finance. *Journal of Business Ethics, 103*(4), 605-619.

Halbritter, G., & Dorfleitner, G. (2015). The wages of social responsibility — Where are they? A critical review of ESG investing. *Review of Financial Economics, 26*, 25-35.

Hall, P. (1992). *The bootstrap and edgeworth expansion*, Springer, New York.

Heimann, M., Pouget, S., Mullet, É., & Bonnefon, J. (2011). The experimental approach to trust in socially responsible investment funds. *Finance and sustainability: Towards a new paradigm? A post-crisis agenda*, 169-183. Emerald Group Publishing Limited.

Henke, H. M. (2016). The effect of social screening on bond mutual fund performance. *Journal of Banking & Finance, 67*, 69-84.

Hörisch, J., Ortas, E., Schaltegger, S., & Álvarez, I. (2015). Environmental effects of sustainability management tools: An empirical analysis of large companies. *Ecological Economics, 120*, 241-249.

Humphrey, J. E., Lee, D. D., & Shen, Y. (2012). Does it cost to be sustainable? *Journal of Corporate Finance, 18*(3), 626-639.

Humphrey, J. E., Warren, G. J., & Boon, J. (2016). What is different about socially responsible funds? A holdings-based analysis. *Journal of Business Ethics, 138*(2), 263-277.

Jagannathan, R., & Wang, Z. (1996). The conditional CAPM and the cross-section of expected returns. *The Journal of Finance, 51*(1), 3-53.

Javed, M., Rashid, M. A., & Hussain, G. (2016). When does it pay to be good–A contingency perspective on corporate social and financial performance: Would it work? *Journal of Cleaner Production, 133*, 1062-1073.

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance, 48*(1), 65-91.

Jobson, J. D., & Korkie, B. M. (1981). Performance hypothesis testing with the Sharpe and Treynor measures. *The Journal of Finance, 36*(4), 889-908.

Kempf, A., & Osthoff, P. (2007). The effect of socially responsible investing on portfolio performance. *European Financial Management, 13*(5), 908-922.

Lahiri, S. (2003). *Resampling methods for dependent data*, Springer, New York.

Ledoit, O., & Wolf, M. (2008). Robust performance hypothesis testing with the Sharpe ratio. *Journal of Empirical Finance, 15*(5), 850-859.

Lee, J., Yen, P., & Chan, K. C. (2013). Market states and disposition effect: Evidence from Taiwan mutual fund investors. *Applied Economics, 45*(10), 1331-1342.

Leggio, K. B., & Lien, D. (2003). An empirical examination of the effectiveness of dollar-cost averaging using downside risk performance measures. *Journal of Economics and Finance, 27*(2), 211-223.

Leite, P., & Cortez, M. C. (2015). Performance of European socially responsible funds during market crises: Evidence from France. *International Review of Financial Analysis, 40*, 132-141.

Loh, W. (1987). Calibrating confidence coefficients. *Journal of the American Statistical Association, 82*(397), 155-162.

Loughran, T., & Ritter, J. R. (2000). Uniformly least powerful tests of market efficiency. *Journal of Financial Economics, 55*(3), 361-389.

Lu, W., Chau, K., Wang, H., & Pan, W. (2014). A decade's debate on the nexus between corporate social and corporate financial performance: A critical review of empirical studies 2002–2011. *Journal of Cleaner Production, 79*, 195-206.

Luo, C., Seco, L., & Wu, L. B. (2015). Portfolio optimization in hedge funds by OGARCH and Markov switching model. *Omega, 57*, 34-39.

Lydenberg, S., & White, A. (2015). Responsible investment indexes: Origins, nature and purpose. In T. Hebb, J. P. Hawley, A. G. Hoepner, A. L. Neher & D. Wood (Eds.), *The Routledge handbook of responsible investment*, 1st ed., 527-535. Routledge, New York.

Margolis, J. D., & Walsh, J. P. (2003). Misery loves companies: Rethinking social initiatives by business. *Administrative science quarterly*, *48*(2), 268-305.

Margolis, J. D., Elfenbein, H. A., & Walsh, J. P. (2009). Does it pay to be good... and does it matter? A meta-analysis of the relationship between corporate social and financial performance.

Meligkotsidou, L., Vrontos, I. D., & Vrontos, S. D. (2009). Quantile regression analysis of hedge fund strategies. *Journal of Empirical Finance, 16*(2), 264-279.

Memmel, C. (2003). Performance hypothesis testing with the Sharpe ratio. *Finance Letters, 1*, 21-23.

Mollet, J. C., & Ziegler, A. (2014). Socially responsible investing and stock performance: New empirical evidence for the US and European stock markets. *Review of Financial Economics, 23*(4), 208-216.

Neher, A. L., & Hebb, T. (2015). The responsible investment atlas–an introduction. *The Routledge Handbook of Responsible Investment*, 53-57. Routledge, New York.

Newey, W., & West, K. (1987). A Simple Positive Semi-definite, Heteroskedasticity and Autocorrelation consistent Covariance Matrix. *Econometrica*, *55*(3), 703–708.

Nicolosi, M., Grassi, S., & Stanghellini, E. (2014). Item response models to measure corporate social responsibility. *Applied Financial Economics*, *24*(22-24), 1449-1464.

Nilsson, J. (2008). Investment with a conscience: Examining the impact of pro-social attitudes and perceived financial performance on socially responsible investment behavior. *Journal of Business Ethics, 83*(2), 307-325.

Nilsson, J. (2015). Stakeholders of responsible investment: Retail investors. In T. Hebb, J. P. Hawley, A. G. Hoepner, A. L. Neher & D. Wood (Eds.), *The Routledge handbook of responsible investment*, 1st ed., 485-493. Routledge, New York.

Nofsinger, J., & Varma, A. (2014). Socially responsible funds and market crises. *Journal of Banking & Finance, 48*, 180-193.

Orlitzky, M., Schmidt, F. L., & Rynes, S. L. (2003). Corporate social and financial performance: A meta-analysis. *Organization Studies, 24*(3), 403-441.

Osthoff, P. (2015). What matters to SRI investors? In T. Hebb, J. P. Hawley, A. G. Hoepner, A. L. Neher & D. Wood (Eds.), *The Routledge handbook of responsible investment*, 1st ed., 705-724. Routledge, New York.

Pagan, A. R., & Sossounov, K. A. (2003). A simple framework for analysing bull and bear markets. *Journal of Applied Econometrics, 18*(1), 23-46.

Politis, D. N., & Romano, J. P. (1992). A circular block-resampling procedure for stationary data. *Exploring the Limits of Bootstrap*, 263-270.

Renneboog, L., Ter Horst, J., & Zhang, C. (2008). The price of ethics and stakeholder governance: The performance of socially responsible mutual funds. *Journal of Corporate Finance, 14*(3), 302-322.

Revelli, C., & Viviani, J. (2015). Financial performance of socially responsible investing (SRI): What have we learned? A meta-analysis. *Business Ethics: A European Review, 24*(2), 158-185.

Scholtens, B. (2008). A note on the interaction between corporate social responsibility and financial performance. *Ecological Economics, 68*(1), 46-55.

Scholtens, B. (2015). Contemporary issues in responsible finance and investment. In T. Hebb, J. P. Hawley, A. G. Hoepner, A. L. Neher & D. Wood (Eds.), *The Routledge handbook of responsible investment*, 1st ed., 575-592. Routledge, New York.

Scholz, H. (2007). Refinements to the Sharpe ratio: Comparing alternatives for bear markets. *Journal of Asset Management, 7*(5), 347-357.
Schröder, M. (2007). Is there a difference? the performance characteristics of SRI equity indices. *Journal of Business Finance & Accounting, 34*(1-2), 331-348.
Sharpe, W. F. (1966). Mutual fund performance. *The Journal of Business, 39*(1), 119-138.
Sortino, F. A., & Price, L. N. (1994). Performance measurement in a downside risk framework. *The Journal of Investing, 3*(3), 59-64.
Sortino, F. A., & Van Der Meer, R. (1991). Downside risk. *The Journal of Portfolio Management, 17*(4), 27-31.
Statman, M. (2006). Socially responsible indexes. *The Journal of Portfolio Management, 32*(3), 100-109.
Statman, M., & Glushkov, D. (2009). The wages of social responsibility. *Financial Analysts Journal*, 65(4), 33-46.
Statman, M., & Glushkov, D. (2016). Classifying and measuring the performance of socially responsible mutual funds. *The Journal of Portfolio Management, 42*(2), 140-151.
Utz, S., & Wimmer, M. (2014). Are they any good at all? A financial and ethical analysis of socially responsible mutual funds. *Journal of Asset Management, 15*(1), 72-82.
Van de Velde, E. Vermeir, W. & Corten, F. (2005). Corporate social responsibility and financial performance. *Corporate Governance: The International Journal of Business in Society, 5*(3), 129-138.
Van Duuren, E., Plantinga, A., & Scholtens, B. (2016). ESG integration and the investment management process: Fundamental investing reinvented. *Journal of Business Ethics, 138*(3), 525-533.
Van Hoorn, A. (2015). The global financial crisis and the values of professionals in finance: An empirical analysis. *Journal of Business Ethics, 130*(2), 253.
Wimmer, M. (2013). ESG-persistence in socially responsible mutual funds. *Journal of Management and Sustainability, 3*(1), 9-15.

Chapter 4: The performance of socially responsible stock portfolios: international evidence

Guillermo Badía[1], María Ceu Cortez[2], and Luis Ferruz[1]

[1]Faculty of Economics and Business, University of Zaragoza,
Zaragoza 50005, Spain.

[2]NIPE - School of Economics and Management, University of Minho,
Gualtar, 4710-057, Braga, Portugal.

Abstract

This paper investigates the financial performance of international stock portfolios formed according to Environment, Social and Governance criteria. Using a robust multi-factor model, we find that European portfolios with the highest Social scores perform better than their low-socially rated counterparts, whereas those that score high on the Governance dimension tend to underperform their low-ranked peers. In Asia Pacific, firms with high Governance scores show some outperformance. In other regions, there are no significant differences between high- and low-ranked portfolios on ESG criteria. We further find that the impact of social screening on portfolio performance is market state and geographically dependent.

Keywords Socially responsible investing; Corporate social responsibility; Social screening processes; Portfolio performance evaluation; Market states

This study has been accepted for presentation in the EFMA International Congress and also in the PhD Consortium of the Foro de Finanzas.
Currently, this study is under review in the International Review of Economics and Finance, journal indexed in the Journal of Citation Report (JCR).

4.1 Introduction

In recent decades, investment management has undergone a progressive adaption process in which conventional financial objectives are increasingly being complemented by non-financial attributes such as environment, social and governance (ESG) criteria. This trend reflects an increasing awareness of environmental, social, and ethical issues that is strongly influencing the purchase decisions of investors (Mollet and Ziegler, 2014). Socially responsible investment (SRI) appeals to investors who wish to go beyond the financial utility of their investments and also derive non-financial utility from holding securities that reflect their social values (Auer, 2016; Auer and Schuhmacher, 2016). Additionally, ESG issues are becoming an important part of investors' decision-making process by helping them to identify firms' long-term opportunities and risks. According to the 2016 Global Sustainable Investment Review, in 2016 there were $22.89 trillion of assets being professionally managed under responsible investment strategies globally, representing an increase of 25 percent since 2014. In 2016, 53% of the total professional managed assets in Europe used SRI strategies, whereas the proportion of SRI relative to total managed assets in the US represented 22%. And in Australia/New Zealand, 51% of assets under professional management were directed to socially responsible investing.

The basic idea of SRI is to apply a set of screens to the available investment universe, in order to select or exclude assets based on ESG criteria (Auer, 2016). In practice, there is a range of SRI strategies, such as integration, positive/best-in-class screening, ethical/negative screening, governance and engagement, etc. All of these aim to drive funds towards socially responsible firms with constructive sustainable projects and policies. Extant studies indicate that not all socially responsible investors are alike, and screens are an important instrument for distinguishing socially responsible practices that serve specific segments of socially conscious investors (Derwall et al., 2011). From an investors' perspective, the critical issue is whether socially responsible stock selection leads to gains or losses in terms of financial performance. On the firms' side, the question is whether spending resources on corporate social responsibility (CSR) practices will render benefits for the firm and increase its value. If doing good is indeed linked to doing well, firms may be led to behave in a more sustainable way. A positive relationship between social and financial performance would even legitimize CSR on economic grounds (Margolis et al. 2009). There are many empirical studies on the financial consequences of including non-financial criteria in the portfolio selection

process. An important stream of the literature has focused on the financial performance of SRI mutual funds. In general, these studies find that there are no significant differences between the performance of SRI mutual funds and conventional funds.[15] However, assessing the financial impact of SRI by evaluating the performance of actively managed SRI mutual funds has some shortcomings. For instance, as Brammer et al. (2006), and Kempf and Osthoff (2007) point out, there are confounding effects - such as fund manager skills and management fees - that may make it difficult to identify the performance that is due to the social characteristics of the underlying holdings. Furthermore, the fact that a mutual fund is classified as a SRI fund does not assure investors that they truly hold stocks of socially responsible companies, thereby suggesting that the label 'socially responsible' may be more of a marketing strategy used by the fund industry. In fact, Utz and Wimmer (2014) show that, on average, SRI funds do not hold more ethical firms than conventional funds, and Statman and Glushkov (2016) even find evidence of 'closet' SRI funds, which are conventional funds that avoid investing in unethical stocks. To overcome the limitations associated to studies on actively managed SRI mutual funds, an alternative approach to evaluate the financial effects of SRI involves evaluating the performance of synthetic portfolios formed on firms' social characteristics. This paper follows this approach to evaluating socially responsible investments.

The purpose of this paper is to investigate the financial performance of international stock portfolios based on CSR criteria. We form portfolios of stocks with high and low sustainability scores and investigate the performance of such portfolios using multi-factor models. Sustainability is measured by an aggregate measure of CSR as well as three indicators of its individual dimensions: Environment, Social and Corporate Governance. Out database comprises international companies covered by ASSET4 ESG database between 2002 and 2017. Previous studies that address the performance of socially screened synthetic portfolios suffer from some limitations and inconsistencies, namely, (1) the majority of prior evidence only refers to the US and European stock markets; (2) with the exception of Badía et al. (2017), previous studies do not compare the performance of SRI portfolios of different regions worldwide; (3) there are studies that measure CSR through one of its individual dimension only, whereas others consider an aggregate construct of CSR; (4) most studies do not evaluate the influence of specific

[15]For a review of studies on the performance of SRI equity funds see, for instance, Capelle-Blancard and Monjon (2012), and Revelli and Viviani (2015).

industries on the financial performance of SRI stock portfolios; (5) in several studies assessing European firms, undersized samples are used; (6) up-to-date evidence is lacking; and (7) additionally, some researchers who document that SRI stock portfolios outperform conventional investments investigate whether there could be a 'time effect', i.e., whether SRI returns were better in earlier years and yet declined in more recent periods. Consistent with the error-in-expectations hypothesis, superior financial performance linked to SRI in earlier times can be a result of a mispricing that disappeared once markets learned how to price these stocks correctly (Derwall et al., 2011), and, hence, markets have adjusted to a pricing equilibrium. The findings of Derwall et al. (2011), Borgers et al. (2013), and Halbritter and Dorfleitner (2015) support this argument by documenting a notable downward movement of abnormal returns of SRI portfolios over time. However, Kempf and Osthoff (2007), Statman and Glushkov (2009), and Mollet et al. (2013) find no significant differences in SRI portfolio performance between sub-periods. These inconsistent results suggest that splitting the sample merely into sub-periods may provide a cursory interpretation of the behavior of SRI portfolio performance in time. In a different perspective, recent studies have provided evidence that socially responsible investments perform differently according to the state of the market, (e.g., recession and expansion periods). Examples of such studies include Nofsinger and Varma (2014), Becchetti et al. (2015), and Leite and Cortez (2015) on SRI equity funds; Henke (2016) on SRI fixed-income funds; and Brzeszczynski and McIntosh (2014), Carvalho and Areal (2016), and Badía et al. (2017) on SRI stock portfolios. We suggest that the inconsistent results of prior studies dividing the sample period in sub-samples may have neglected an important effect, specifically, the impact of different market states.

Hence, our main contributions to the existing literature are fivefold: (1) we extend the analysis on the impact of including socially responsible screens on investment portfolios performance to additional geographical areas (North America, Europe, Japan, and Asia Pacific); (2) we compare the financial performance of SRI portfolios of these regions to each other; (3) we form portfolios based on an aggregate measure of CSR as well as on three of its specific ESG dimensions; (4) we evaluate the influence of specific industries on the financial performance of SRI stock portfolios; and finally, (5) we assess the financial performance of SRI stock portfolios over different market states: bear, bull and mixed market periods. Considering the growth of socially responsible investments in international capital markets and the intensifying global competition, the valuation

implications of sustainability in an international context is of practical interest to management, investors and regulators worldwide.

The remainder of the paper is organized as follows: Section 2 discusses the financial effects of SRI, providing an overview of the most influential studies related to the financial performance of SRI stock portfolios and discussing their limitations. Section 3 describes the data. Section 4 presents and discusses the empirical methodology and results, and section 5 summarizes the main results and presents some concluding remarks.

4.2 The financial effects of SRI

4.2.1 Theoretical arguments

There are two contrasting hypothesis on the effects of socially responsible investing in portfolio financial performance. The underperformance hypothesis is consistent with a traditional view of CSR that suggests a negative link between CSR and corporate financial performance (CFP). According to this perspective, supported by Friedman (1970), integrating environmental and social aspects in firm policies will have negative financial implications, since it implies internalizing additional costs. As Eccles et al. (2014) mention, high-sustainability firms may underperform since, for instance, they may discard valuable business opportunities that do not match their policies and values, or they may experience higher labor costs by providing more benefits to their employees. A further argument supporting the underperformance of SRI portfolios stems directly from portfolio theory, that sustains that portfolios formed on the basis of a limited set of investment opportunities will not be mean-variance efficient. Additionally, the screening process implies increased monitoring and information costs that also penalize financial performance (Cortez et al., 2009). Finally, there is evidence that stocks shunned by socially responsible investors (e.g., tobacco, alcohol and weapons) yield abnormal returns (Hong and Kacperzyck, 2009; Statman and Glushkov, 2009; Derwall et al., 2011). Since socially responsible investors typically avoid these stocks, they will not be able to benefit from those returns to the extent conventional investors do so.

Nevertheless, proponents of SRI claim that socially screened investing may result in a higher financial performance. This argument is supported by many empirical studies

that document a positive relation between CSR and CFP and valuation.[16] The outperformance hypothesis is consistent with stakeholder theory (Freeman, 1984) and the argument that integrating stakeholders' interests creates value for shareholders (Jensen, 2001). For instance, responsible firms may outperform by constituting confident supply chains, by innovating and developing products that maintain environmental constraints, and by attracting and retaining high-quality human capital (Eccles et al., 2014). Preston and O'Bannon (1997) also argue that satisfying the interests of different corporate stakeholders enhances a firm's reputation, resulting in a positive impact on its financial performance. They note that, since CSR involves constantly assessing corporate influences and relationships with stakeholders and the environment, it allows management to recognize and react to evolving strategic opportunities and challenges. In this line of reasoning, the use of social screens can help investors identify companies with better management skills (Bollen, 2007), and consequently benefit from an improved financial performance.

4.2.2 A critical look at prior empirical evidence

This section provides an overview of the most influential studies related to the financial performance of SRI stock portfolios. Table 1 summarizes empirical studies that assess the financial performance of SRI stock portfolios and the links to the seven controversial issues outlined in the introduction.[17]

According to the column 'Portfolio Construction', most studies form a portfolio with high-sustainability firms on the basis of a CSR indicator (high-ranked) and another with low-sustainability firms (low-ranked), and compare their financial performance by forming a differences portfolio, obtained by subtracting the low-ranked portfolio returns from the returns of the high-ranked portfolio (H-L analysis). Other studies (e.g., Filbeck et al., 2009; Edmans, 2011; Mollet et al., 2013; Brzeszczynski and McIntosh, 2014; Auer, 2016; Badía et al., 2017) compare the performance of portfolios of high-sustainability stocks to conventional benchmarks (CCB analysis).

[16]For a more in-depth discussion of the empirical studies in the field, see for example, the review studies of Margolis and Walsh (2003), Orlitzky et al. (2003), Margolis et al. (2009), Lu et al. (2014), and Javed et al. (2016).

[17]We do not include in this discussion studies analyzing the relationship between reputation and financial performance (as reputation is a more vague concept, not so easily measured as the other components of ESG) nor those that do not use risk-adjusted measures to evaluate portfolio performance.

Table 4-1. Prior evidence of studies assessing the financial performance of SRI stock portfolios

This table presents the most influential studies that form stock portfolios on the basis of social criteria and evaluate their financial performance. Column *Market state* identifies the studies assessing the impact of different market states on the financial performance of SRI stock portfolios. Column *Specific market* shows the market that the authors analyse: US (the United States), EU (European Union), UK (the United Kingdom), and AP (Asia Pacific). Column *Start* shows the first year analyzed by the authors and column *End* identifies the last year analyzed. Column *Industry effect* indicates the studies assessing the influence of specific industries on the financial performance of SRI stock portfolios. Column *Individual or an aggregate score* shows the criterion used regarding the ESG dimensions to assess the effect of CSR on the financial performance. SD means that authors use a singular dimension; MD means that authors evaluate several dimensions (multi-dimension) and (XD), indicates the authors focus on X individual dimensions; AD means that authors use an aggregate dimension. Column *Results: Statistic financial differences?* shows whether there are statistically significant differences between the financial performance of compared portfolios (for instance, high minus low-ranked stocks). Column *Portfolio construction* shows the procedure to compare the financial performance between portfolios. H-L indicates that authors form a high and a low portfolio and assess financial performance differences; CCB indicates that authors compare high sustainable firms to conventional benchmarks. (1*) a portfolio of firms that score high on employee relations and a sin stocks portfolio is formed. (2*) depending on the ESG dimension used, investors in Europe tend to pay a price for socially responsible investing.

Article	Market state	Specific market	Start	End	Industry effect	Individual or an aggregate score	Results: Statistic financial differences?	Portfolio construction
Filbeck and Preece (2003)	Not	US	1987	1999	Not	SD - Employee	Yes +	CCB
Derwall et al. (2005)	Not	US	1995	2003	Yes	SD - Environment	Yes +	H-L
Van de Velde et al. (2005)	Not	EU	2000	2003	Not	MD (5D) & AD	Not	H-L
Brammer et al. (2006)	Not	UK	2002	2004	Yes	MD (3D) & AD	Yes -	CCB
Kempf and Osthoff (2007)	Not	US	1992	2004	Not	MD (6D) & AD	Yes +	H-L
Galema et al. (2008)	Not	US	1992	2006	Yes	MD (6D)	Not	H-L
Brammer et al. (2009)	Not	US	2000	2004	Yes	AD	Not	CCB
Filbeck et al. (2009)	Not	US	2000	2007	Not	AD	Yes +	CCB
Statman and Glushkov (2009)	Not	US	1997	2007	Not	MD (7D) & AD	Yes +	H-L
Derwall et al. (2011)	Not	US	1992	2008	Not	SD - Employee	Not analyzed	H & L (1*)
Edmans et al. (2011)	Not	US	1984	2009	Yes	SD - Employee	Yes +	CCB
Humphrey et al. (2012)	Not	UK	2002	2010	Yes	AD	Not	H-L & CCB
Borgers et al. (2013)	Not	US	1992	2009	Not	AD	Yes +	H-L
Lee et al. (2013)	Not	US	1998	2007	Yes	AD	Not	H-L
Mollet et al. (2013)	Not	EU	2002	2009	Yes	AD	Yes +	CCB
Brzeszczynski and McIntosh (2014)	Yes	UK	2000	2010	Not	AD	Not	CCB
Eccles et al. (2014)	Not	US	1993	2010	Yes	AD	Yes +	H-L
Mollet and Ziegler (2014)	Not	US & EU	1998	2009	Not	AD	Not	H-L
Halbritter and Dorfleitner (2015)	Not	US	1991	2012	Not	MD (ESG) & AD	Not	H-L
Auer (2016)	Not	EU	2004	2012	Not	MD (ESG) & AD	Yes +	CCB
Auer and Schuhmacher (2016)	Not	US & EU & AP	2004	2012	Yes	MD (ESG)	US & AP Not / EU Yes -/+ (2*)	H-L & CCB
Carvalho and Areal (2016)	Yes	US	1998	2010	Not	SD - Employee	Not	CCB
Badía et al. (2017)	Yes	US & EU & AP	2005	2014	Not	AD	Yes +	CCB

Table 1 confirms that the majority of prior studies address the US and the EU markets. This could be justified given the noteworthy proportions of assets that are professionally managed under responsible investment strategies in these countries. Auer and Schuhmacher (2016), and Badía et al. (2017) are the exception, since they extend their scope to Asia-Pacific countries in a multiregional analysis, and evaluate the financial performance of firms from the US, European, and Asia-Pacific markets. And although Auer and Schuhmacher (2016) compare, within each region, high- and low-ranked stock portfolios, they do not evaluate the relative financial performance of each regional portfolio. Badía et al. (2017) compare the returns of regional portfolios to each other from a retail investor's perspective and find outperformance of SRI portfolios in some specific geographical areas. Considering this evidence as well as the heterogeneity in the patterns of development of SRI across countries (Neher and Hebb, 2015), SRI financial performance should be further documented and compared in different regions. The extension of SRI research to other geographical areas is further motivated by Hörisch et al. (2015), who indicate that country-specific factors tend to affect the relationship between corporate social and financial performance. Additionally, investors' ESG concerns can also differ from region to region. For instance, Eccles et al. (2011) find that European investors are more concerned with environmental information, while US investors are more interested in governance issues. In turn, Cortez et al. (2012) identify geographical differences in the investment style of socially responsible funds. Furthermore, given the progressive saturation of the SRI market in the US (Mollet et al., 2013), SRI diffusion and expansion in other regions could be indicative of a productive niche for positive abnormal returns.

The information in the column 'Individual or an aggregate score' of Table 1 shows the criteria used to measure CSR. Authors such as Filbeck and Preece (2003), Derwall et al. (2005), Derwall et al. (2011), Edmans (2011), and Carvalho and Areal (2016) focus their attention on a singular dimension of CSR: environment or employee relations. While this type of analysis shows the impact of a specific dimension of CSR on financial performance, it is restrictive to draw general conclusions about the effect of general features of sustainability on performance. On the other hand, the use of individual dimensions of CSR may be important because relevant characteristics of companies might end up diluted when using a combined measure of CSR (Hoepner et al., 2016). Some authors have used both specific dimensions of CSR as well as an aggregate construct. For instance, within the US market, Kempf and Osthoff (2007)

analyse six different CSR dimensions, together with an aggregate score, and Statman and Glushkov (2009) consider seven dimensions and an aggregate score. However, Galema et al. (2008) leave the overall score aside, while Borgers et al. (2013) only consider an aggregate score. Regarding the European SRI market, studies such as Van de Velde et al. (2005) and Auer (2016) scatter sustainability among different dimensions, while Humphrey et al. (2012) and Mollet and Ziegler (2014) combined measure of CSR. Auer and Schuhmacher (2016), and Badía et al. (2017), who also evaluate the Asia Pacific region, follow different approaches. The former use both an aggregate score and individual dimensions (ESG), whereas the latter only uses an aggregate score. As we discuss below, there are various advantages and disadvantages on using an aggregate or individual dimensions of CSR to qualify the social responsibility of firms. In this paper, we consider both an aggregate measure of CSR as well as measures of its individual components (Environment, Social, and Governance), which allows us to recognise the individual influence of each singular dimension, along with the effect of an overview score on portfolio financial performance.

The column 'Industry effect' of Table 1 shows the studies assessing the influence of specific industries on the financial performance of SRI stock portfolios. Several studies such as Eccles et al. (2014) and Auer and Schuhmacher (2016) analyse the industry effects in socially responsible investing, whereas Mollet and Ziegler (2014), Halbritter and Dorfleitner (2015), Auer (2016), and Badía et al. (2017) do not look at this aspect. Focusing on the US market, Derwall et al. (2005), Galema et al. (2008), Edmans (2011), Lee et al. (2013), and Eccles et al. (2014) evaluate specific-industry influences, but Filbeck and Preece (2003), Kempf and Osthoff (2007), Statman and Glushkov (2009), Borgers et al. (2013), and Carvalho and Areal (2016) do not. A similar scenario is observed in European and multiregional studies. These ambiguous findings are surprising since some studies (e.g., Derwall et al. 2005; Brammer et al. 2006; Porter and Kramer, 2006; Hoepner et al., 2010) have shown that different industries differ in terms of the concrete CSR opportunities and risks, and that these may influence the relationship between CSR and CFP. In this vein, we investigate the industry-sensitivity of SRI stock portfolios.

An additional limitation related to prior evidence is the under-sized sample bias stressed by Auer (2016). We confirm this evidence in such studies as Van de Velde et al. (2005), and Brammer et al. (2006). Table 1 shows that empirical evidence is just documented up to 2014. The column 'End' of Table 1 displays the last year analyzed by prior

studies. We can see that the more up-to-date sample period (to 2014) is studied by Badía et al. (2017). As noted in reports such as the Global Sustainable Investment Review of both 2014 and 2016, SRI expansion has been intensive in recent periods. Therefore, we emphasise that more contemporary evidence is required on the financial influence of considering SRI aspects.

The most controversial issue associated with SRI is the financial impact of social screening. Observing Table 1, we note that the results are inconclusive. The column 'Results: Statistic financial differences?' concerns the financial implications of SRI. While some studies do not find significant financial differences between high- and low-sustainable firms, or conventional benchmarks (e.g., Van de Velde et al., 2005; Galema et al., 2008; Brammer et al., 2009; Lee et al. 2013), others support the positive financial performance of SRI (e.g., Derwall et al., 2005; Kempf and Osthoff, 2007; Edmans, 2011; Eccles et al., 2014; and Badía et al., 2017). In contrast, Brammer et al. (2006) and Auer and Schuhmacher (2016) find evidence of negative performance in some European countries. The information presented in Table 1 shows that accounting for SRI aspects in the portfolio selection process tends to have no negative effects on financial performance in the majority of cases.

Finally, the column 'Market state' identifies the studies assessing the impact of different market states on the financial performance of SRI stock portfolios. As noted previously, the recent literature has documented a significant effect of different market phases on the performance of SRI investment funds, indices, portfolios, etc. However, Table 1 shows that, with the exception of Brzeszczynski and McIntosh (2014), Carvalho and Areal (2016), and Badía et al. (2017), no prior studies of SRI stock portfolios have distinguished SRI performance in different market states. Brzeszczynski and McIntosh (2014) identify bull and bear periods via the Woodward and Anderson (2009) approach and find that there are no financial differences in performance between bull and bear markets. However, they simply observe raw return differences, without testing for statistical differences in alphas. By a more sophisticated methodology - specifically, through a conditional model that allows both risk and performance to vary over different market phases - Carvalho and Areal (2016) find that both the financial performance and the systematic risk of a SRI stock portfolio remain unaffected in bear markets. They use the Pagan and Sossounov (2003) procedure to identify bull and bear periods. In a similar process, Badía et al. (2017) document that SRI portfolios outperform conventional investments during bull periods and abide neutral during bear

markets. As mentioned previously, there are some studies that divide the sample period into sub-periods, but they could render only a cursory review of the performance evolution. Consequently, we analyse the financial performance of SRI stock portfolios in different market states (bull and bear markets).

In sum, this review discusses the limitations and shortcomings of prior empirical studies. In this paper, we aim to overcome these limitations in the evaluation of SRI stock portfolio performance.

4.3 Data

We assess the financial consequences of social screening processes on a global scope. To form portfolios, we use the social responsibility ratings of companies provided by Thomson Reuters ASSET4 ESG database.[18] The ASSET4 ESG rating classifies stocks based on roughly 700 individual data points, then combined into over 250 key performance indicators (KPIs), and later aggregated into a framework of 18 categories to form the four ESG pillars (Economic, Environmental, Social and Corporate Governance pillars). As part of the calculation rating method, all companies are measured against the complete firm universe. The ASSET4 ESG database further computes an overall ESG score that includes the four pillars mentioned above. Since we wish to form portfolios on the basis on non-economic indicators, we do not use the overall ESG score computed by the database. Instead, we compute a combined ESG score as an equally-weighted average of these three individual scores, as in Auer (2016). Instead, we construct an overall ESG score as an equally-weighted average score of the three pillars: Environment, Social and Corporate Governance.

We analyze an international sample including firms from 23 countries over the period January 2002 to December 2017. In order to mitigate a potential short country-specific sample bias that could reduce the power of our tests, we combine the 23 countries into four diversified regional portfolios: North America (NA), that includes the United States and Canada; Europe (EU), that includes Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, and the United Kingdom; Japan (JA); and Asia Pacific (AP), that includes Australia, New Zealand, Hong Kong, and Singapore. We follow the allocation of Fama and French (2012, 2017) who group countries in regions mainly by geographic

[18]The Thomson Reuters ASSET4 ESG database has been used in prior studies evaluating the financial performance of SRI stock portfolios (e.g., Eccles et al., 2014; Halbritter and Dorfleitner, 2015).

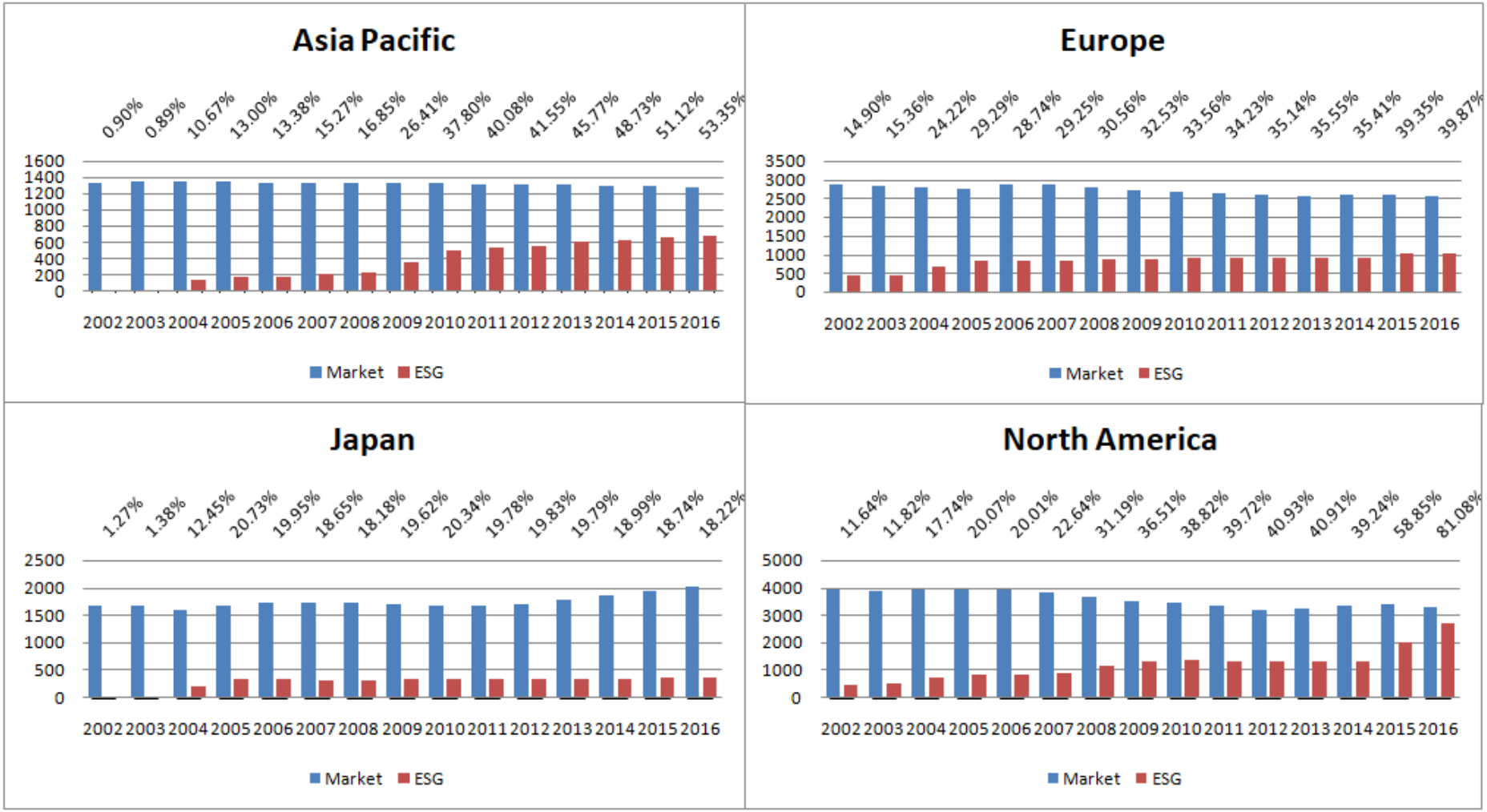
location and market integration. Monthly discrete returns of all stocks are computed based on the total return series (in US dollars) collected from the Thomson Reuters database. In line with Cooper et al., (2004) and Asem (2009), in order to minimize nontrading and microstructure-induced biases, stocks whose prices are below $1 at the beginning of the holding period and those with a stable price for two consecutive months are screened out. Survivorship bias does not affect our results since we use the full ASSET4 universe, thus including both active and inactive stocks.

Since we investigate the financial performance of SRI in different regions, it is interesting to show the proportion of firms evaluated in each region relative to the local markets. To this purpose, we track the local stock exchange where stocks are traded and evaluate the percentage of firms with ESG qualifications on them. For instance, for the JP market, the TOPIX index is considered as the local market since stocks with ESG information in ASSET4 are included in this index. A year-by-year analysis is done on the constituents of the index. Then, we calculate the percentage of stocks with ESG values provided by ASSET4 on the local stock exchange market.[19]

Figure 1 shows that the proportion of stocks with ESG scores on each region has increased progressively over the sample period, with exception of the JP market that shows a notable growth of stocks in the TOPIX index just in recent periods. Nonetheless, the number of stocks with ESG ratings in this market has increased, and with the exception of the two first years, around 20% of firms have social ratings. A similar picture is documented in the AP market for the two first years, although the evolution in this region is somewhat different. It is striking that at the beginning of the sample period only around 1% of firms have ESG scores and, yet, in the two most recent periods, more than a half of the firms are rated. As for the EU market, firms with ESG scores have continuously increased across the sample period, representing around

[19]Indices for the EU market are: ATX, BEL 20, OMX COPENHAGEN, SBF 120, FTSE All-Share, FTSE MIB, ATHEX COMPOSITE, OMX HELSINKI, IRELAND SE OVERALL, MADRID SE GENERAL, AMSTERDAM (AEX), OMX AFFARSVARLDENS GENERAL, OSLO SE OBX, PSI GENERAL, DAX 30 PERFORMANCE, SWISS ALL SH; for the NA market are: S&P 500 COMPOSITE, S&P/TSX COMPOSITE INDEX, NASDAQ COMPOSITE; for the AP market are: ASX ALL ORDINARIES, HANG SENG, NZX Main Board, STRAITS TIMES INDEX; and for the JP market is: TOPIX.

Figure 4-1. Proportion of stocks with ESG scores on each region over time (2002-2016)



40% of stocks on local markets in the last year. As expected, these figures allow us to recognize that the firms in the NA market are the most rated ones. It is also interesting to note the notable increase of socially rated firms after the beginning of the international financial crisis. Finally, it is also worth mentioning the substantial growth in the proportion of firms with ESG scores in the NA market in the most recent periods.

In sum, we can observe that number of firms that are rated according to their ESG concerns has increased progressively, surely reflecting an increase in the market-investor demand for this kind of information. Since investors are the main user of that information, it represents an additional evidence of the growing interest for knowing extra-financial information (ESG) of firms.

The use of aggregate or individual dimensions of CSR to qualify for the social responsibility of firms is a debatable issue. A first approach suggests that an analysis of specific dimensions of CSR is likely to be important, since different aspects may have differential impacts, depending on the nature of the firm's business (Van de Velde et al., 2005). As Galema et al. (2008) stress, aggregated CSR measures may confound relationships among different aspects of CSR and CFP. Initiatives such as using energy-saving technology may reduce operating costs, but practices like flexible scheduling may enhance productivity and reduce absenteeism, which may in turn make it easier to recruit and retain outstanding staff (Brammer et al., 2006). As Hoepner et al. (2016) note, important social features may be hidden by using an aggregated measures of CSR. Another viewpoint argues that for many investors a firm overall CSR indicator is more useful than an indicator that reflects an individual dimension of CSR (Boutin-Dufresne and Savaria, 2004). As Lee et al. (2013) point out, most investors do not include only environmental, social, or governance criteria in their decisions. Moreover, Wimmer (2013) highlights that not all investors have a deep understanding of what exactly SRI entails. Consequently, offering an overall CSR measure helps investors to select SRI stocks. Arguably, the first approach is more closely related to a firm view, and the second is more likely to be associated with the investor's perspective. In any case, we consider both an aggregate CSR score and individual dimensions of its components: the Environment, Social, and Governance performance, enabling us to assess not only the influence of CSR on financial performance, but also the individual influence of each of its individual dimension.

Figures 2-4 show the evolution of the mean values ESG and the aggregate scores for each region. We observe that under the most demanding cut-off level (10%), both high-

ranked firms and low-ranked firms of different regions score close to each other in terms of the Environment and Social dimensions. However, some remarkable differences appear in terms of the scores associated to the Governance dimension, in particular the low Governance scores of JP high-ranked firms which are even lower than the scores of low-ranked NA firms. This is not surprising considering the concerns related to transparency, independence, auditing and monitoring functions of JP firms. Despite some initiatives to improve the governance of JP firms, such as the Corporate Governance Code of 2015 (JSIF, 2017), the Governance scores of these firms are still far behind those of the rest of regions. These findings support our decision of evaluating JP firms separately from AP firms. Otherwise, our results for AP firms may be affected by the low scores of JP firms. Auer and Schuhmacher (2016) evaluate AP firms including JP firms in this region and find that AP firms rate higher in the Governance than in Environment and Social criteria. However, our data uncovers the fact that JP firms rate lower in Governance than in Social and Environmental aspects. On the Governance criteria, and whatever cut-off used, we find that NA high-ranked firms stand out, although EU high-ranked firms, and even AP high-ranked firms, score close to them in recent periods. Furthermore, NA low-ranked firms achieve really high Governance scores compared to other regions. This finding suggests that NA firms are particularly concerned with governance issues, and it is in line with Halbritter and Dorfleitner (2015) who find that US firms also rate highest in governance aspects. Figures 2-4 also show that EU firms, both high- and low-ranked, have the highest Social values over the sample period. This result suggests that EU firms are more concerned on paying further attention to social demands than firms of other regions. According to the Environment dimension, both EU and JP firms excel, reaching the highest scores over the sample period. In general, these observations suggest that ESG concerns across regions are different. This is in line with several studies that document regional and cultural idiosyncrasies in socially responsible investing. Louche and Lydenberg (2006), for instance, exploring the development and practices of SRI in the US and EU markets, show that there are some differences in terms of motivations and screening strategies implemented. They note that the emphasis placed on environment is stronger in Europe than in the US. The environment was at the forefront of the European implementation of SRI. The divergent patterns of SRI are reflected on the higher Environmental scores of

Figure 4-2. Mean ESG scores of low-ranked (10% cut-off) and high-ranked firms (90% cut-off)

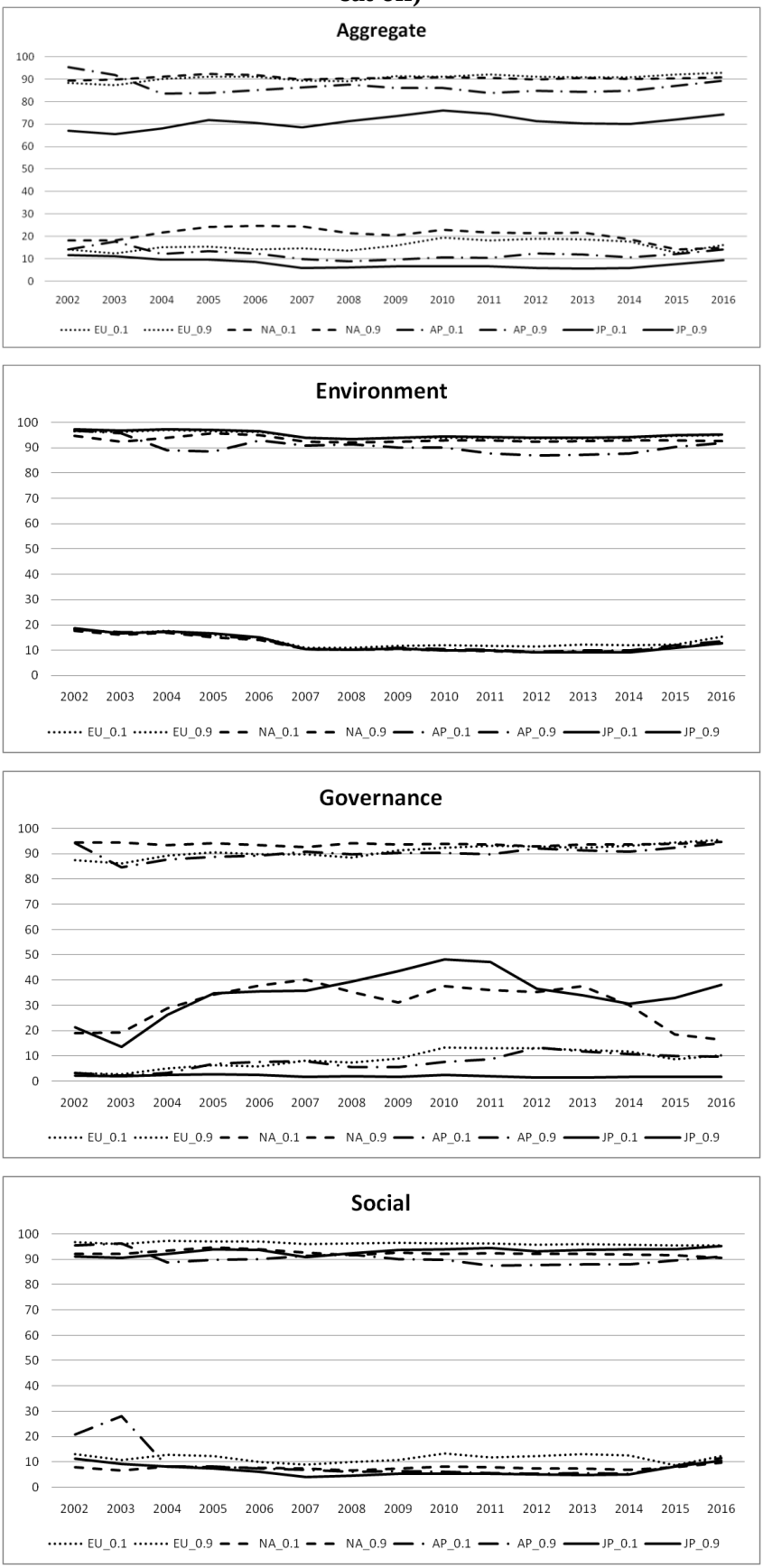
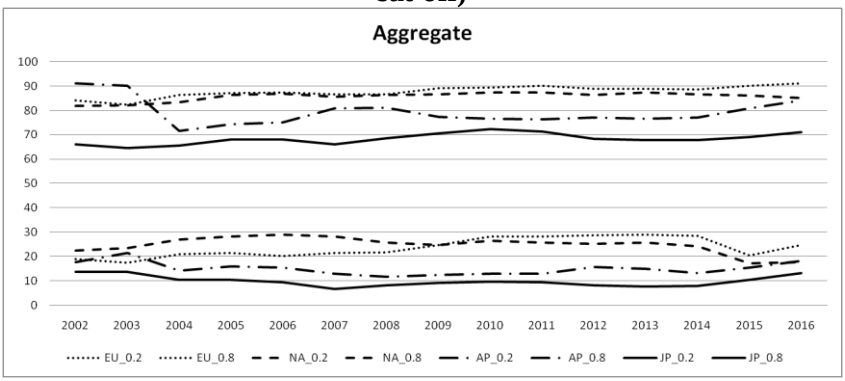
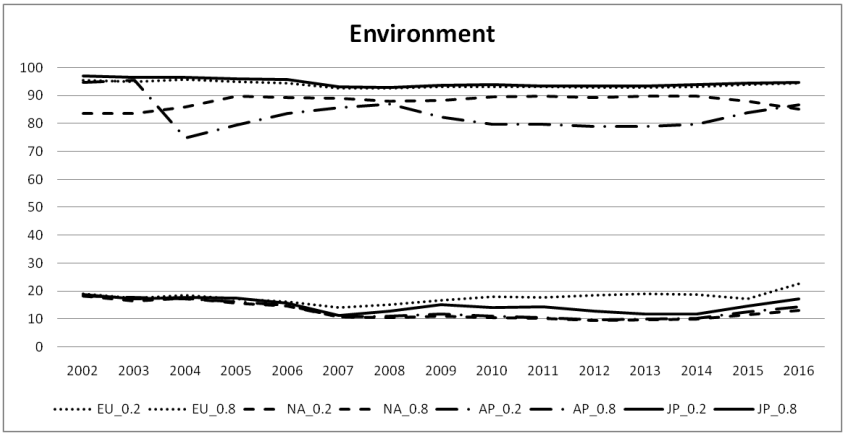
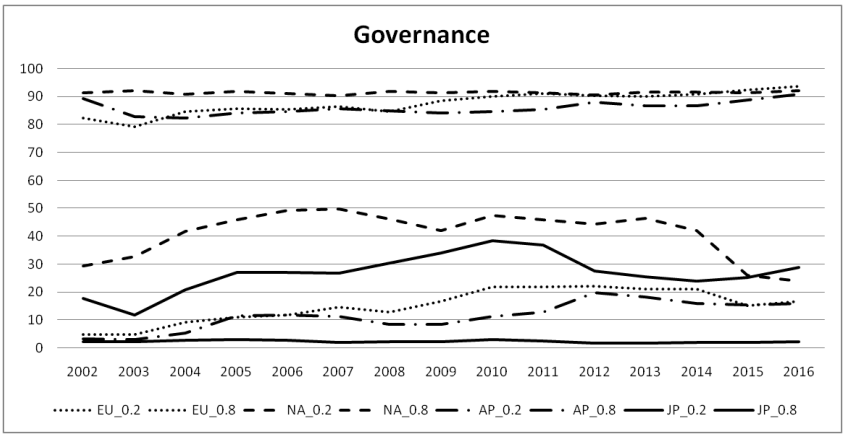


Figure 4-3. Mean ESG scores of low-ranked (20% cut-off) and high-ranked firms (80% cut-off)







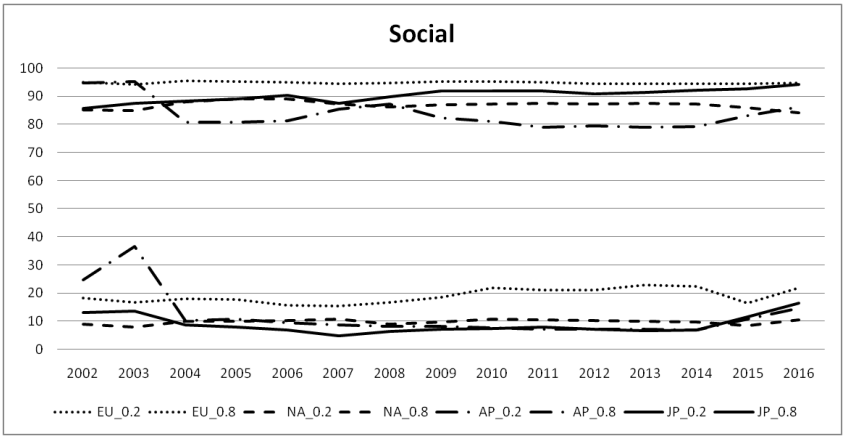
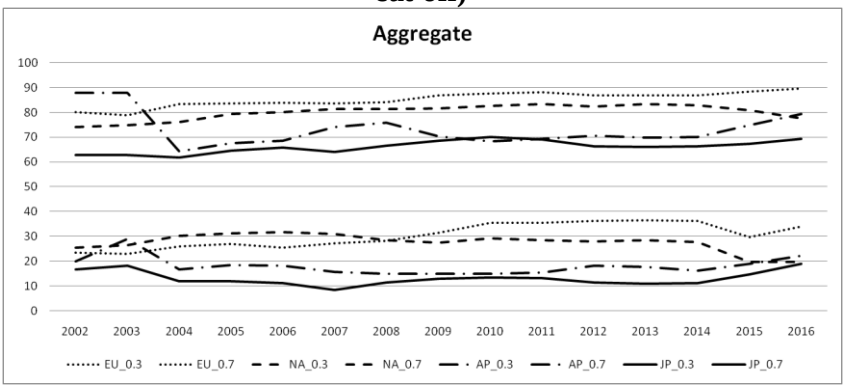
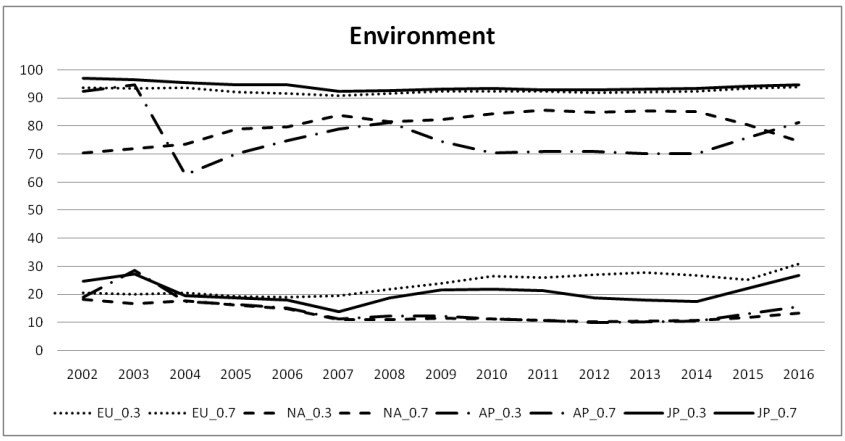
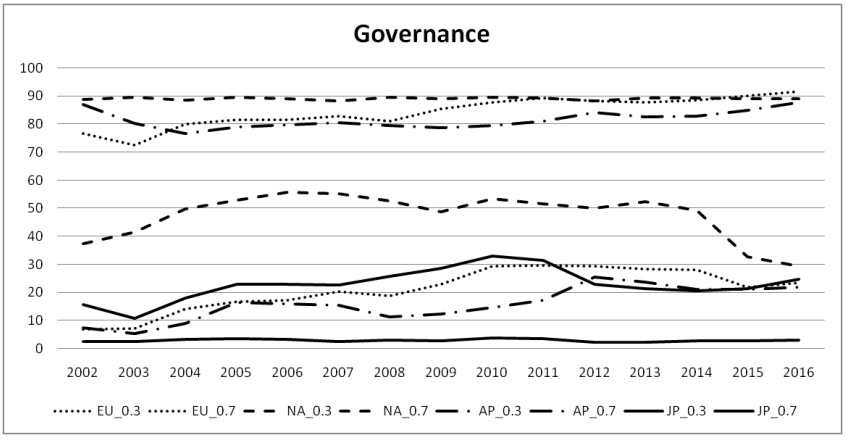
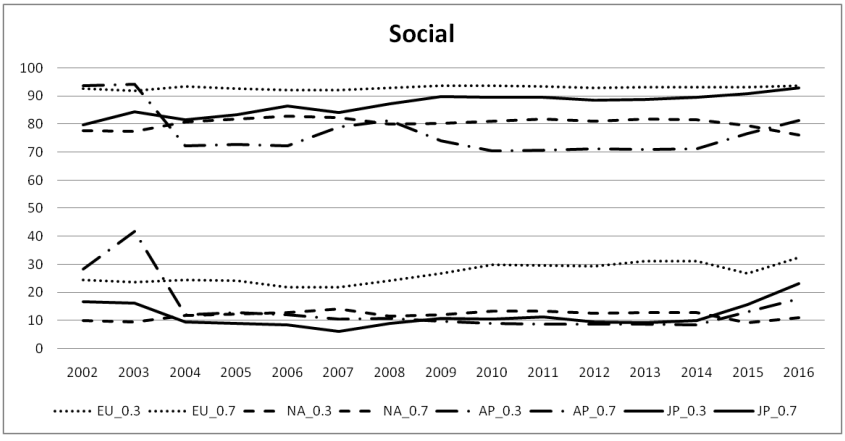


Figure 4-4. Mean ESG scores of low-ranked (30% cut-off) and high-ranked firms (70% cut-off)









EU firms compared to NA firms. In turn, Sakuma and Louche (2008) examine the emergence and development of SRI in Japan and show that it was also particularly linked to an eco-dimension (for instance, eco-funds were the starters in this market). They note indeed that SRI in Japan holds several similarities with Europe in relation to shared SRI purposes, actors, and strategies. Our evidence is in this vein concerning the Environment dimension since we find that JP firms and EU firms have a close evolution and similar high values over the sample period.

4.4 Empirical implementation and results

4.4.1 Portfolio formation

To test the financial consequences of considering social screens in the investment process, each year we form equally-weighted portfolios of stocks of companies based on their social ratings in the previous year. The high-rated portfolio comprises stocks with the best socially rated companies and the low-rated portfolio includes those with the worst socially rated companies. As in prior studies (e.g., Van de Velde et al., 2005; Kempf and Osthoff, 2007; Derwall et al., 2011; Halbritter and Dorfleitner, 2015; and Auer, 2016), we use different cut-offs to form the portfolios (10%, 20%, and 30%), thus allowing us to evaluate portfolios that are more restricted or more broad with respect to the social criteria used. Portfolios are formed for each ESG dimension and for the aggregate score. Then, we form the difference portfolio, which is obtained by subtracting the low-ranked portfolio returns from the returns on the high-ranked stock portfolio, thus representing a strategy of going long in the high-rated stocks and short in the low-rated stocks. The analysis of the performance of the long-short portfolios enables us to conclude whether there are statistically significant differences between the performance of high- and low-rated portfolios. To assess performance differences of SRI firms among regions, we compare the regional high-ranked portfolios to each other. Table 2 presents descriptive statistics of the ESG regional portfolios under different cut-offs. Although in most cases high-rated portfolios yield a higher average return than low-rated portfolios, the differences between average returns are not statistically significant whatever region, ESG dimension, and cut-off level considered. We can also see that inside each region average returns do not change substantially for portfolios formed on different cut-offs. For instance, in EU, the mean return of high-rated portfolios across the sample period is almost the same at the 20% and 30% cut-off levels considering the aggregate score. A similar picture is observed regarding the low-

rated portfolios of AP at the 10% and 20% cut-off levels (Environment score), and the high-rated portfolios of JP under the 10% and 20% cut-off levels (Social score). The level of portfolios' standard deviations allows us to observe that the higher returns of high-rated portfolios are generated together with large volatility. In some cases, differences in standard deviations are even significant. These findings encourage the use of risk-adjusted measures to evaluate financial performance. Finally, as in the case of average returns, we can observe a similar pattern in standard deviation differences among portfolios within regions and different cut-offs.

Table 4-2. Descriptive statistics of regional portfolios

This table presents a summary statistics of high (H) and low (L) ESG regional portfolios. Mean (SD) is the average month return (standard deviation) of portfolios over the sample period: from January 2002 to December 2017. Portfolios at the 10% [10], 20% [20], and 30% [30] cut-off level for the Aggregate (AD), Environment (ENV), Social (SOC), and Corporate Governance (CGV) dimensions are formed from stocks in North America (NA), Europe (EU), Asia Pacific (AP), and Japan (JP).The asterisks are used to represent the statistically significant differences at the 1% (***), 5% (**) and 10% (*) significance levels on tests of equality of mean and variances (t-test and F-test) between high and low portfolios.

Mean	AD			ENV			SOC			CGV		
	H	L		H	L		H	L		H	L	
NA [10]	0.01342	0.01121		0.01228	0.01189		0.01142	0.01113		0.01324	0.01118	
NA [20]	0.01201	0.01199		0.01193	0.01203		0.01090	0.01137		0.01250	0.01199	
NA [30]	0.01198	0.01226		0.01116	0.01250		0.01143	0.01177		0.01250	0.01187	
EU [10]	0.01261	0.01016		0.01395	0.01122		0.01421	0.01043		0.01180	0.01099	
EU [20]	0.01366	0.01029		0.01313	0.01118		0.01458	0.01018		0.01278	0.01055	
EU [30]	0.01367	0.01063		0.01348	0.01116		0.01404	0.01052		0.01296	0.01138	
AP [10]	0.01580	0.01283		0.01662	0.01266		0.01337	0.01339		0.01654	0.01123	
AP [20]	0.01538	0.01195		0.01507	0.01276		0.01421	0.01306		0.01477	0.01165	
AP [30]	0.01430	0.01214		0.01510	0.01208		0.01381	0.01280		0.01447	0.01222	
JP [10]	0.00946	0.00660		0.01137	0.01011		0.01011	0.00930		0.00817	0.00878	
JP [20]	0.00954	0.00878		0.00946	0.00920		0.01013	0.00815		0.00906	0.00932	
JP [30]	0.01007	0.00880		0.00994	0.00947		0.00992	0.00858		0.00887	0.00982	
SD	AD			ENV			SOC			CGV		
	H	L		H	L		H	L		H	L	
NA [10]	0.05374	0.04553	**	0.05114	0.04672		0.05662	0.04264	***	0.05271	0.04959	
NA [20]	0.05300	0.04516	**	0.05044	0.04737		0.05555	0.04456	***	0.05252	0.05005	
NA [30]	0.05349	0.04695	*	0.05198	0.04793		0.05385	0.04616	**	0.05243	0.04988	
EU [10]	0.06584	0.05923		0.06717	0.06215		0.06659	0.05893		0.06410	0.06086	
EU [20]	0.06452	0.05965		0.06418	0.06247		0.06531	0.05950		0.06389	0.05989	
EU [30]	0.06380	0.05982		0.06369	0.06159		0.06446	0.05973		0.06327	0.06041	
AP [10]	0.07143	0.06152	**	0.06582	0.06287		0.07346	0.05906	***	0.06960	0.06580	
AP [20]	0.06684	0.06100		0.06399	0.05920		0.07118	0.06101	**	0.06654	0.06490	
AP [30]	0.06675	0.06165		0.06693	0.06120		0.07109	0.06136	**	0.06574	0.06600	
JP [10]	0.04869	0.04944		0.04888	0.05095		0.04994	0.04656		0.04246	0.05046	**
JP [20]	0.04559	0.04620		0.04369	0.04904		0.04930	0.04487		0.04127	0.04721	*
JP [30]	0.04307	0.04525		0.04319	0.04751		0.04637	0.04455		0.04132	0.04702	*

4.4.2 Financial performance

To evaluate portfolio performance, we compute alphas from a multi-factor model, as for example in Van de Velde et al. (2005), Edmans (2011), Humphrey et al. (2012), and

Badía et al. (2017). These studies examine performance using the four-factor Carhart (1997) model that captures the risk premiums associated with size and value versus growth (as in Fama and French, 1993) as well as momentum, (motivated by Jegadeesh and Titman,1993). More recently, Fama and French (2015) identify an additional set of risk factors in the US market. They test a five-factor asset pricing model that adds the profitability and investment factors to the market, size, and value-growth factors. Their results show that the inclusion of these new risk factors to the Fama and French (1993) three-factor model improves the capacity to explain the cross-section of expected stock returns. Fama and French (2017) test the five-factor model specification in an international context (North America, Europe, Japan, and Asia Pacific), and also find satisfactory results. In spite of the fact that these additional risk factors may capture relevant sources of systematic risk, none of the prior studies on the performance of SRI portfolios uses them. We follow Fama and French (2018) and use a six-factor model that includes the five factors of the Fama and French (2015) five-factor model augmented by the momentum factor. The model is estimated given the following equation:

$$R_{it} - R_{Ft} = a_i + b_i Mkt_t + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + m_i MOM_t + e_{it}$$

(Eq. 4-1)

where R_{it} is the dollar return on portfolio *i* for month *t*, R_{Ft} is the risk-free rate (the one-month US Treasury bill rate), Mkt_t is the value-weighted market portfolio return minus the risk-free rate. The remaining variables are the differences between the returns on diversified portfolios of small and large stocks (SMB_t), high and low B/M stocks (HML_t), stocks with robust and weak profitability (RMW_t), stocks of low and high investment firms, conservative minus aggressive, (CMA_t), and winning and losing stocks in the past year (MOM_t). e_{it} is a zero-mean residual. α_i is the estimated financial performance measure of the portfolio, and b_i, s_i, h_i, r_i, c_i, and m_i represent the estimated risk measures associated with the different risk factors. The independent variables are obtained from Professor Kenneth French's website.

Since the relationship between SRI and financial performance may be affected by industry characteristics in terms of the specific ESG opportunities and exposure (Derwall et al., 2005; Brammer et al., 2006), and considering the widely-held view among investors that industry-specific ESG criteria provide useful information

(Humphrey et al., 2012),[20] we investigate the industry-adjusted portfolio performance following Geczy et al. (2003). We extend the multi-factor model (eq.1) to include controls for industry biases. To this end, for each region, we first run a regression of the 25 TRBC[21] industry indices on the market index, thus making sure that they are orthogonal to the market. A new 'cleaned' index is created by the sum of the intercept and the residuals of the regression. The cleaned industry index is then only capturing industry specific return characteristics. Next, a principal components analysis is performed to drive industry factors. These factors are added to equation (1) to control for industry effects that are not captured, as follows:

$$R_{it} - R_{Ft} = a_i + b_i Mkt_t + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + m_i MOM_t + \sum_{k=1}^{\gamma} l_k IP_{it} + e_{it}$$

(Eq. 4-2)

where $\sum_{k=1}^{\gamma} l_k IP_{it}$ represents the γ principal component factors capturing industry effects on portfolio returns. γ principal components are selected for regions:[22] for the NA portfolio, we use six industry components; for the EU portfolio, four industry components; for the AP portfolio, five industry components; and for the JP portfolio, four industry components. Previous studies use an alternative number of components depending on the market evaluated (see, for example, Derwall et al., 2005; Humphrey et al., 2012).

Panel A of Table 3 displays, for each region and ESG dimension, the alphas of the long-short portfolios under the different cut-offs.[23] In most cases, there are no statistical significant differences in the performance of portfolios of firms with higher social scores and those with lower social scores. We find positive effects of sustainable screening processes on the Governance dimension in the AP market, on the Social dimension in the EU market, and on the Environment dimension in AP markets[24]. We further observe in EU a negative effect of screening at the Governance dimension, since some EU high-rated portfolios underperform low-rated ones. We note (Figures 2-4) that

[20]In fact, DiBartolomeo and Kurtz (1999), Porter and Kramer (2006), and Hoepner et al. (2010) find evidence that industry exposures drive the financial performance of SRI portfolios.
[21]The Thomson Reuters Business Classification.
[22]Principal components with eigenvalues superior to 1 are selected for each region.
[23]As we are focusing on the performance of SRI portfolios, only the alphas of the long-short portfolios are reported. Nonetheless, coefficients related to specific beta risk-factors are available upon request.
[24]The analysis focuses on the results that are statistically significant at least at the 5% level.

EU high-rated firms achieved the highest Governance scores in recent periods, close to NA and AP portfolios. Still, some AP high-rated portfolios do outperform their low-rated peers. This contrasting effect is in line with Eccles et al. (2011), and Cortez et al. (2012) who identify that the concerns and investment styles of investors are different across regions, and also with Louche and Lydenberg (2006), and Neher and Hebb (2015) who suggest that regional and cultural idiosyncrasies aspects may affect financial performance of socially responsible investing. Furthermore, the existence of abnormal returns from CSR depends not only on the existence of a positive effect of these activities on the firms' valuation, but mostly on whether financial markets reflect all of the value-relevant information that relates to CSR practices. If investors do not fully understand how to adequately assess CSR and its impact on the fundamental value of the company, there may be opportunities for them to yield higher abnormal returns from socially responsible investing (Derwall et al., 2011). Our results regarding the Governance dimension suggest that investors in different regions are in different stages with regards to their understanding of the impact of good corporate governance practices on firms' valuation. Our results further show that screening processes based on the Social dimension have a strong positive effect on EU firms. High-ranked firms on the Social dimension outperform their low-ranked counterparts whatever the cut-off considered. This effect, associated to the high Social scores of EU firms (Figures2-4), suggests that European markets are rewarding the Social performance of firms. Regarding the Environment dimension, we find that these screens only have a positive effect on AP firms under the less demanding SRI level (30% cut-off). However, after industry-adjustments (Panel B), this effect disappears.

These findings are of interest to SRI investors since they suggest they can form portfolios that are consistent with their beliefs and personal values without being negatively affected in terms of financial performance. Our results support those of Auer and Schuhmacher (2016) for the AP, and NA market, but contrast with those that the authors obtain regarding the specific ESG dimensions that negatively affect the financial performance of EU high-rated portfolios. The findings of Mollet and Ziegler (2014) are also in line with our results for NA, while contrasting with those of the EU market. Yet, it is important to keep in mind that Mollet and Ziegler (2014) only measure sustainability with an aggregate indicator of CSR. In fact, as previously mentioned, any comparison of results with those of previous studies (summarized in Table 1) must be

done with caution, as some of them assess CSR by using measures of its individual dimensions and others use an aggregate measure of CSR score.

Table 4-3. Regional portfolio performance

This table shows estimates of alpha (abnormal returns) of the long-short portfolios at the 10% [10], 20% [20], and 30% [30] cut-off level for the Aggregate dimension (AD), Environment (ENV), Social (SOC), and Corporate Governance (CGV). The long-short portfolio is formed by subtracting the returns of the high-ranked portfolio from the returns of the low-ranked portfolio (H-L).Portfolios are formed for North America (NA), Europe (EU), Asia Pacific (AP), and Japan (JP).Panel A displays results of estimating portfolio financial performance by means of the Fama and French (2015) five-factor model augmented by the momentum factor (eq.1). The independent variables are obtained from Professor Kenneth French's website. Panel B shows the results of extending equation (1) to control for industry effects following the approach of Geczy et al. (2003). The multi-factor models are estimated by OLS based on the heteroskedasticity and autocorrelation adjusted errors of Newey and West (1987). One-month US T-bills proxy for the risk-free rate. The asterisks are used to represent the statistically significant coefficients at the 1% (***), 5% (**) and 10% (*) significance levels. The full sample period is from January 2002 to December 2017.

Panel A: Alphas of the long-short portfolios: H-L analysis

NA	AD	ENV	SOC	CGV
Long-Short [10]	0.0023	0.0017	0.0001	0.0025*
Long-Short [20]	-0.0001	0.0011	-0.0008	0.0008
Long-Short [30]	-0.0002	-0.0004	-0.0003	0.0007
EU	AD	ENV	SOC	CGV
Long-Short [10]	-0.0004	0.0026	0.0037**	-0.0038**
Long-Short [20]	0.0022*	0.0018	0.0049***	-0.0011
Long-Short [30]	0.0018	0.0025*	0.0035***	-0.0017
AP	AD	ENV	SOC	CGV
Long-Short [10]	0.0036	0.0044	-0.0006	0.0062**
Long-Short [20]	0.0042*	0.0023	0.0013	0.0037*
Long-Short [30]	0.0028	0.0034**	0.0015	0.0024
JP	AD	ENV	SOC	CGV
Long-Short [10]	0.0013	0.0017	-0.0002	-0.0009
Long-Short [20]	-0.0004	0.0006	0.0002	-0.0004
Long-Short [30]	0.0009	0.0005	0.0002	-0.0009

Panel B: Alphas of the long-short portfolios controlling for industry effects: H-L analysis

NA	AD	ENV	SOC	CGV
Long-Short [10]	0.0023*	0.0018	0.0002	0.0025*
Long-Short [20]	0.0000	0.0011	-0.0007	0.0008
Long-Short [30]	-0.0001	-0.0003	-0.0003	0.0007
EU	AD	ENV	SOC	CGV
Long-Short [10]	-0.0004	0.0026*	0.0037**	-0.0037**
Long-Short [20]	0.0022*	0.0017	0.0049***	-0.0010
Long-Short [30]	0.0018	0.0025*	0.0034***	-0.0016
AP	AD	ENV	SOC	CGV
Long-Short [10]	0.0032	0.0049*	-0.0007	0.0064**
Long-Short [20]	0.0040*	0.0021	0.0012	0.0036*
Long-Short [30]	0.0027	0.0032*	0.0015	0.0024
JP	AD	ENV	SOC	CGV
Long-Short [10]	0.0012	0.0017	-0.0004	-0.0009
Long-Short [20]	-0.0003	0.0007	0.0002	-0.0004
Long-Short [30]	0.0010	0.0006	0.0002	-0.0009

Panel B of Table 3 shows that portfolio performance results after industry-adjustments are similar to those of Panel A. Humphrey et al. (2012) do not find that industry-specific

criteria affects the financial performance of UK SRI portfolios formed on an aggregate measure of CSR. The findings of Mollet et al. (2013), who also use an aggregate indicator of CSR for the EU market, are similar. Our results on the portfolios formed on the aggregate ESG score are in line with them. On the other hand, whereas Derwall et al. (2005), focusing on the Environmental performance of US firms, find that the difference in financial performance between high- and low-rated portfolios increases when industry effects are considered, we do not find significant industry effects in that market. Our results for the NA market are thus in line with Galema et al. (2008), and Lee et al. (2013), who observe that industry components do not have a significant effect on the financial performance of portfolios based on ESG criteria.

Table 4 shows the results on the relative financial performance across regional portfolios. Panel A displays, for pairs of regions, the alphas of the long-short portfolios under different cut-offs. The results show that, in general, high-rated SRI regional portfolios do not show statistical significant differences in performance. The exception refers to high-rated NA portfolios formed on the Governance dimension and a 10% cut-off, which outperform JP portfolios of high-rated firms. This evidence cannot be disassociated from the high scores of NA firms and the low scores of JP firms on the Governance dimension (Figures 2-4). It suggests that screening processes based on the Governance dimension can affect the portfolio financial performance across regions. However, in general, our results suggest that regional-specific aspects do not seem to have a significant effect on the financial performance of SRI firms when they are compared to each other. The financial performance of high-rated portfolios is similar across regions. These findings together the previous ones (Table 3) indicate that significant financial performance differences should be sought within regions. In this regard, despite the patterns of development of SRI not being homogenous across countries (Neher and Hebb, 2015), and a roughly progressive saturation of SRI in markets such as the US (Mollet et al., 2013), our evidence does not uncover differences in the financial consequences of investin in high-ranked firms across regions. In fact, our results suggest that high socially rated firms across regions share similar benefits from specific characteristics in terms of risk and opportunities.

Panel B of Table 4 shows the results controlling for industry effects. The relative performance after industry-adjustment is mostly akin to our previous results. The only difference is observed for JP&NA in terms of the Governance dimensions, indicating that the performance of JP high-rated firms is negatively affected by industry effects.

Table 4-4. Relative financial performance of regional portfolios

This table shows estimates of alpha (abnormal returns) of the long-short portfolios at the 10% [10], 20% [20], and 30% [30] cut-off level for the Aggregate dimension (AD), Environment (ENV), Social (SOC), and Corporate Governance (CGV). The long-short portfolio is formed by subtracting the high-ranked portfolio returns of a region from the returns on the high-ranked portfolio of another one (H-H). Panel A shows results of estimating portfolio financial performance by means of the Fama and French (2015) five-factor model augmented by the momentum factor (eq.1). The independent variables are obtained from Professor Kenneth French's website. Global factors are used to estimate the financial portfolio performance among regions. Panel B shows results of extending equation (1) to control for industry effects following the approach of Geczy et al. (2003). 25 TRBC industry global indices are used and principal components with eigenvalues superior to 1 are selected. The multi-factor models are estimated by OLS based on the heteroskedasticity and autocorrelation adjusted errors of Newey and West (1987). One-month US T-bills proxy for the risk-free rate. The asterisks are used to represent the statistically significant coefficients at the 1% (***), 5% (**) and 10% (*) significance levels. The full sample period is from January 2002 to December 2017.

Panel A. Alphas of the long-short portfolios: H-H analysis				
EU & NA	AD	ENV	SOC	CGV
Long-Short [10]	-0.0045	-0.0034	-0.0005	-0.0050*
Long-Short [20]	-0.0009	-0.0020	0.0016	-0.0030
Long-Short [30]	-0.0008	-0.0006	0.0001	-0.0029
AP & NA	AD	ENV	SOC	CGV
Long-Short [10]	0.0000	-0.0004	-0.0016	0.0012
Long-Short [20]	0.0021	0.0000	-0.0001	0.0013
Long-Short [30]	0.0007	0.0006	-0.0012	0.0006
JP & NA	AD	ENV	SOC	CGV
Long-Short [10]	-0.0059*	-0.0052	-0.0042	-0.0055**
Long-Short [20]	-0.0041	-0.0051	-0.0030	-0.0037
Long-Short [30]	-0.0033	-0.0035	-0.0035	-0.0044
AP & EU	AD	ENV	SOC	CGV
Long-Short [10]	0.0045	0.0030	-0.0012	0.0062*
Long-Short [20]	0.0030	0.0020	-0.0017	0.0043
Long-Short [30]	0.0015	0.0011	-0.0014	0.0034
EU & JP	AD	ENV	SOC	CGV
Long-Short [10]	0.0014	0.0018	0.0038	0.0005
Long-Short [20]	0.0032	0.0031	0.0046	0.0008
Long-Short [30]	0.0025	0.0029	0.0036	0.0015
AP & JP	AD	ENV	SOC	CGV
Long-Short [10]	0.0059	0.0048	0.0026	0.0067
Long-Short [20]	0.0062	0.0051	0.0029	0.0051
Long-Short [30]	0.0039	0.0041	0.0022	0.0049
Panel B: Alphas of the long-short portfolios controlling for industry effects: H-H analysis				
EU & NA	AD	ENV	SOC	CGV
Long-Short [10]	-0.0045*	-0.0034	-0.0005	-0.0050*
Long-Short [20]	-0.0010	-0.0020	0.0015	-0.0030
Long-Short [30]	-0.0009	-0.0006	0.0001	-0.0029
AP & NA	AD	ENV	SOC	CGV
Long-Short [10]	0.0000	-0.0004	-0.0017	0.0012
Long-Short [20]	0.0020	-0.0001	-0.0002	0.0013
Long-Short [30]	0.0006	0.0005	-0.0013	0.0005
JP & NA	AD	ENV	SOC	CGV
Long-Short [10]	-0.0060	-0.0053	-0.0043	-0.0055*
Long-Short [20]	-0.0042	-0.0052	-0.0031	-0.0038
Long-Short [30]	-0.0033	-0.0035	-0.0035	-0.0044
AP & EU	AD	ENV	SOC	CGV
Long-Short [10]	0.0045	0.0030	-0.0012	0.0062*
Long-Short [20]	0.0030	0.0020	-0.0017	0.0043
Long-Short [30]	0.0014	0.0011	-0.0014	0.0034
EU & JP	AD	ENV	SOC	CGV
Long-Short [10]	0.0014	0.0019	0.0038	0.0005
Long-Short [20]	0.0032	0.0031	0.0046	0.0008
Long-Short [30]	0.0025	0.0029	0.0036	0.0015

(Continuation)

AP & JP	AD	ENV	SOC	CGV
Long-Short [10]	0.0060	0.0048	0.0026	0.0067
Long-Short [20]	0.0062	0.0051	0.0029	0.0051
Long-Short [30]	0.0039	0.0040	0.0022	0.0049

These findings show that, in general, the relative financial performance of SRI firms among regions is not affected when controlling for industry exposure. Our findings are novel in this regard since previous studies do not compare SRI regional portfolios taking into account industry effects.

4.4.3 Financial performance under different market conditions

To analyse the market state effect on financial performance, the first step is to identify market phases across the sample period. To this purpose, we use the Pagan and Sossounov (2003), hereafter PS, approach. PS develop a procedure to identify the peaks and troughs of a stock market index. A peak is established at t time in the case of the event $PK = [\ln P_{t-8}, \dots, \ln P_{t-1} < \ln P_t > \ln P_{t+1}, \dots, \ln P_{t+8}]$ occurs, where P_t represents the quotation of the stock market index, and a trough at time t in the case of the event $TH = [\ln P_{t-8}, \dots, \ln P_{t-1} > \ln P_t < \ln P_{t+1}, \dots, \ln P_{t+8}]$ occurs. Following previous studies (e.g., Leite and Cortez, 2015; Badía et al., 2017), we qualify bear periods as those with a downtrend in the relevant market index of at least 20% from peak to trough. The remaining periods are considered as bull periods. The relevant stock market indices used are: the MSCI North America Index, the MSCI Europe Index, the MSCI Japan Index, and the MSCI Pacific ex Japan Index. Table 5 shows the bear markets identified according to PS (2003).

Table 4-5. Bear market states

This table identifies bear market periods according to the Pagan and Sossounov (2003) procedure. The indices used are the MSCI North America Index, the MSCI Europe Index, the MSCI Japan Index, and the MSCI Pacific ex Japan Index. Consistent with the literature, we require the rise (fall) of the market being greater (less) than either 20%.The window breadth for eight, nine and ten months is evaluated and the same results are obtained.The full sample period is from January 2002 to December 2017.

Portfolio	Start date	Indexvalue (Points)	End date	Indexvalue (Points)	Change in marketindex	Length of bear period (months)
NA	Nov-07	1558.805	Feb-09	776.949	-0.5016	16
EU	Nov-07	2159.770	Feb-09	873.949	-0.5954	16
	May-11	1588.340	May-12	1164.809	-0.2667	13
	Jun-14	1819.889	Feb-16	1391.740	-0.2353	21
AP	Nov-07	1521.787	Feb-09	607.648	-0.6007	16
	Aug-14	1463.360	Feb-16	1025.155	-0.2995	19
JP	Mar-07	3303.140	Feb-09	1720.810	-0.4790	24

The downward trend in prices related to the international financial crisis that emerged in 2007 is identified in all markets. We even observe that the Japanese market somewhat anticipates this crisis (March 2007) compared to other markets. We further identify two additional bear market periods in Europe: from May 2011 to May 2012, and from June 2014 to February 2016. The former can be associated to the Euro sovereign debt crisis, and the latter to the uncertainty about the future of the Greek economy. We also find an additional bear market period in the Asia-Pacific region from August 2014 to February 2016, which can be associated to the slowdown in the growth of the Chinese economy.

Once the market states have been identified, portfolio performance is evaluated by a model that includes two dummy variables, in line with Nofsinger and Varma (2014). This model allows both risk and performance to vary across different market phases, as follows:

$$\begin{aligned} R_{it} - R_{Ft} = {} & a_{Bear}\, D_{Bear,t} + a_{Bull}\, D_{Bull,t} + b_{Bear}\, Mkt_t D_{Bear,t} + b_{Bull}\, Mkt_t D_{Bull,t} \\ & + s_{Bear}\, SMB_t D_{Bear,t} + s_{Bull}\, SMB_t D_{Bull,t} + h_{Bear}\, HML_t D_{Bear,t} \\ & + h_{Bull}\, HML_t D_{Bull,t} + r_{Bear}\, RMW_t D_{Bear,t} + r_{Bull}\, RMW_t D_{Bull,t} \\ & + c_{Bear}\, CMA_t D_{Bear,t} + c_{Bull}\, CMA_t D_{Bull,t} + m_{Bear}\, MOM_t D_{Bear,t} \\ & + m_{Bull}\, MOM_t D_{Bull,t} + e_{it} \end{aligned}$$

(Eq. 4-3)

where $D_{Bear,t}$ is a dummy variable that takes value 1 for bear market periods and zero otherwise, and $D_{Bull,t}$ is a dummy variable that takes value 1 for bull market periods and zero otherwise; α_{Bear} corresponds to the financial performance in bear markets and α_{Bull} in bull markets; b_{Bear}, s_{Bear}, h_{Bear}, r_{Bear}, c_{Bear}, and m_{Bear} correspond to the factor loadings in bear periods; and b_{Bull}, s_{Bull}, h_{Bull}, r_{Bull}, c_{Bull}, and m_{Bull} in bull periods. This specification of the model extends the one used by Nofsinger and Varma (2014) by incorporating the dummy variables both for the alphas and for the risk factors. We are thus the first study to extend the Nofsinger and Varma (2014) approach to the Fama and French (2015) five-factor model augmented by the momentum.

Table 6 displays the alpha estimates of regional portfolios over different market phases. Panel A shows that in the NA market there are no significant differences between high- and low-rated portfolios in both market periods. However, when controlling for industry effects (Panel B), NA high-ranked firms formed on the Environment dimension (30% cut-off) underperform their low-ranked peers in bear periods.

Table 4-6. Financial performance in different market states

This table shows results of estimating alpha (abnormal returns) of the long-short portfolios in different market states at the 10% [10], 20% [20], and 30% [30] cut-off level for the Aggregate dimension (AD), Environment (ENV), Social (SOC), and Corporate Governance (CGV). The long-short portfolio is formed by subtracting the return of the high-ranked portfolio from the returns of the low-ranked portfolio (H-L).Portfolios are formed for North America (NA), Europe (EU), Asia Pacific (AP), and Japan (JP).The Pagan and Sossounov (2003) procedure is used in order to identify different market states (bear and bull).Panel A displays estimates of portfolio financial performance based on equation (3). Panel B shows results of extending equation (3) to control for industry effects following the approach of Geczy et al. (2003). The multi-factor models are estimated by OLS based on the heteroskedasticity and autocorrelation adjusted errors of Newey and West (1987). The asterisks are used to represent the statistically significant coefficients at the 1% (***), 5% (**) and 10% (*) significance levels. The full sample period is from January 2002 to December 2017.

Panel A: Alphas of the long-short portfolios: H-L

	Bear				Bull			
NA	AD	ENV	SOC	CGV	AD	ENV	SOC	CGV
Long-Short [10]	0.0049	0.0081*	0.0000	0.0120*	0.0020	0.0011	-0.0008	0.0012
Long-Short [20]	-0.0053	-0.0009	-0.0026	0.0026	-0.0004	0.0007	-0.0014	0.0006
Long-Short [30]	-0.0026	-0.0052	-0.0015	0.0021	-0.0005	-0.0007	-0.0012	0.0009
EU	AD	ENV	SOC	CGV	AD	ENV	SOC	CGV
Long-Short [10]	0.0008	0.0003	0.0004	0.0011	0.0001	0.0036*	0.0028	-0.0013
Long-Short [20]	-0.0004	-0.0013	0.0007	0.0025	0.0033**	0.0018	0.0038**	0.0001
Long-Short [30]	-0.0002	-0.0024	-0.0009	0.0006	0.0024*	0.0033**	0.0031***	-0.0004
AP	AD	ENV	SOC	CGV	AD	ENV	SOC	CGV
Long-Short [10]	0.0038	0.0000	0.0056	0.0065	0.0029	0.0081**	-0.0001	0.0070**
Long-Short [20]	0.0038	0.0008	0.0016	0.0027	0.0045*	0.0032	0.0018	0.0041
Long-Short [30]	0.0033	0.0031	0.0013	0.0043	0.0028	0.0034*	0.0016	0.0021
JP	AD	ENV	SOC	CGV	AD	ENV	SOC	CGV
Long-Short [10]	-0.0012	-0.0041	0.0019	-0.0094**	0.0003	0.0019	-0.0006	-0.0017
Long-Short [20]	-0.0078**	-0.0079**	-0.0014	-0.0046	-0.0001	0.0015	-0.0001	-0.0009
Long-Short [30]	-0.0078***	-0.0064***	-0.0043***	-0.0052**	0.0012	0.0012	0.0001	-0.0009

Panel B: Alphas of the long-short portfolios controlling for industry effects: H-L

	Bear				Bull			
NA	AD	ENV	SOC	CGV	AD	ENV	SOC	CGV
Long-Short [10]	0.0070	0.0072	0.0006	0.0147*	0.0018	0.0012	-0.0008	0.0010
Long-Short [20]	-0.0060	-0.0044	-0.0032	0.0054	-0.0002	0.0012	-0.0012	0.0003
Long-Short [30]	-0.0036	-0.0087**	-0.0036	0.0042	-0.0004	-0.0002	-0.0009	0.0007
EU	AD	ENV	SOC	CGV	AD	ENV	SOC	CGV
Long-Short [10]	0.0014	-0.0011	0.0001	0.0019	-0.0004	0.0039*	0.0029	-0.0018
Long-Short [20]	-0.0006	-0.0025	-0.0005	0.0037	0.0033**	0.0021	0.0044**	-0.0007
Long-Short [30]	-0.0004	-0.0033	-0.0019	0.0016	0.0023*	0.0035**	0.0034***	-0.0010
AP	AD	ENV	SOC	CGV	AD	ENV	SOC	CGV
Long-Short [10]	0.0077	0.0020	0.0060	0.0149**	0.0017	0.0077*	-0.0001	0.0050
Long-Short [20]	0.0057	0.0022	0.0025	0.0066	0.0040	0.0026	0.0015	0.0031
Long-Short [30]	0.0061	0.0028	0.0049	0.0097**	0.0021	0.0033	0.0007	0.0008
JP	AD	ENV	SOC	CGV	AD	ENV	SOC	CGV
Long-Short [10]	-0.0022	-0.0014	-0.0033	-0.0101	0.0004	0.0016	0.0001	-0.0014
Long-Short [20]	-0.0029	-0.0069	0.0024	-0.0035	-0.0008	0.0014	-0.0007	-0.0010
Long-Short [30]	-0.0053	-0.0046	-0.0024	-0.0039	0.0009	0.0011	-0.0002	-0.0011

In the EU market, there are no significant differences between high- and low-rated portfolios in bear market periods, whereas in bull markets high-rated portfolios based on the Aggregate, Environment, and especially the Social dimension outperform their low-rated counterparts. These effects persist after industry-adjustments (Panel B). These

results are relevant since whereas evaluating the financial performance of EU portfolios across the full sample period only shows significant performance differences on portfolios formed on the Social dimension, doing so over different market states only uncovers performance differentials in bull markets and on several dimensions. Additionally, the underperformance of high-ranked firms on the Governance dimension disappears in both market periods. In the AP market, significant differences also appear in bull markets regarding portfolios formed on the Environment and Governance dimensions. Yet, this outperformance does not survive the adjustment for influential industry characteristics (Panel B). Industry-adjustments are particularly important on the AP market: in bear markets, high-rated portfolios formed on the Governance dimension outperform low-rated ones after industry controls. This evidence suggests that AP well-governed firms show a better resilience in bear periods than those that are inefficiently governed. In the JP market, industry effects are also relevant. We find that during bear markets high-ranked firms underperform low-ranked ones. However, these significant differences disappear after controlling for industry effects. In general, our results are in line with Badía et al. (2017), who also document that the financial performance of SRI portfolios is market state dependent.

Table 4-7. Bear and mixed market periods

This table identifies bear and mixed market periods according to the Pagan and Sossounov (2003) procedure. Mixed market periods are identified when bull and bear periods do not match across markets. The indices used are the MSCI North America Index, the MSCI Europe Index, the MSCI Japan Index, and the MSCI Pacific ex Japan Index. Consistent with the literature, we require the rise (fall) of the market being greater (less) than either 20%.The window breadth for eight, nine and ten months is evaluated and the same results are obtained.The full sample period is from January 2002 to December 2017.

Portfolio	Start date	End date	Period	Bear market	Length of period (months)
EU & NA	Nov-07	Feb-09	Bear	Both	16
	May-11	May-12	Mixed	EU	13
	Jun-14	Feb-16	Mixed	EU	21
AP & NA	Nov-07	Feb-09	Bear	Both	16
	Ago-14	Feb-16	Mixed	AP	19
JP & NA	Mar-07	Oct-07	Mixed	JP	8
	Nov-07	Feb-09	Bear	Both	16
AP & EU	Nov-07	Feb-09	Bear	Both	16
	May-11	May-12	Mixed	EU	13
	Jun-14	Jul-14	Mixed	EU	2
	Ago-14	Feb-16	Bear	Both	19
EU & JP	Mar-07	Oct-07	Mixed	JP	8
	Nov-07	Feb-09	Bear	Both	16
	May-11	May-12	Mixed	EU	13
	Jun-14	Feb-16	Mixed	EU	21
AP & JP	Mar-07	Oct-07	Mixed	JP	8
	Nov-07	Feb-09	Bear	Both	16
	Ago-14	Feb-16	Mixed	AP	19

When identifying market states in different regions, we observe that bull and bear periods do not always match across markets. Table 7 shows that, for instance, whereas the EU market is bearish over the period May 2011 to May 2012, the NA market is not. Likewise, while the AP market is depressed from August 2014 to February 2016, the NA market is not.

Considering this mismatch of economic conditions across different markets, we further analyze portfolio performance in times where a specific market state does not occur simultaneously in matched markets. The 'mixed' market state analysis allows us to identify whether financial performance differences are a result of firms of different regions being affected by different market conditions in opposition to country-specific factors related to SRI. Since our results comparing SRI high-rated firms of different regions show that country-specific factors do not seem to affect the financial performance of SRI stock portfolios, with this procedure we investigate whether significant differences among SRI portfolios of different regions are driven by the different market stages they are experiencing. The findings of Badía et al. (2017) suggest the existence of performance differences among SRI regional portfolios. However, they do not evaluate relative financial differences across regions under different market conditions. Hence, as far as our knowledge, we are the first study in doing so. To evaluate that effect, we extend equation (3) to incorporate a new dummy variable, thereby, $D_{BBear\,,t}$ is a dummy variable that takes value 1 when both markets are over bear periods and zero otherwise, $D_{BBull\,,t}$ is a dummy variable that takes value 1 when both markets are over bull periods and zero otherwise, and $D_{Mixed\,,t}$ is a dummy variable that takes value 1 when a market is over a bull period and the other one is over a bear period, i.e. mixed market states, and zero otherwise. These periods are identified in Table 7.

Table 8 shows estimates of alphas of regional portfolios across different market conditions. In Panel A we observe that SRI portfolios perform similarly when both markets are experiencing bull periods. However, when controlling for industry effect we observe performance differences between JP and NA firms on the Governance dimension in bull periods. When both markets are in bear periods, we find several portfolio performance differences. AP firms formed on the Governance dimension outperform NA firms, although these differences disappear after industry-adjustments. We also find that EU firms formed on the Governance dimension outperform NA firms.

These results suggest that EU firms with high Governance scores tend to have an increased resilience to crisis compared to NA firms. But the most relevant results are observed in mixed market states. Strong significant performance differences are observed between JP and NA firms as well as between AP and EU firms. The fact that EU firms are experiencing two bear market periods when AP firms are in bull periods seems to lead to significant performance differences among the portfolios. As for the relative performance between JP and NA firms, although JP firms are suffering an additional bear period in comparison to NA firms, JP firms outperform NA firms during the mixed market state. In spite of the fact that the NA market started the bear period later (Nov-2007), during June and July 2007, the NA market went down 5% whereas the JP market went down only the 0.5%. We also find that, in contrast to bear market periods, NA firms outperform EU firms in terms of the Governance dimension in mixed periods. Furthermore, we find that after controls for industries (Panel B) AP firms perform better than JP firms in mixed periods on several portfolios. These results highlight how different market conditions affect the performance of SRI portfolios.

In sum, these findings contrast to those comparing SRI firms of different regions over the full sample period (Table 4). Previously, we found that high-ranked firms of different regions perform similarly. However, this new analysis shows significant differences among SRI portfolios of different regions as a result of firms of different regions being influenced by distinct economic conditions.

Table 4-8. Relative financial performance across regional portfolios in different market states

This table shows results of estimating alpha (abnormal returns) of the long-short portfolios in different market states at the 10% [10], 20% [20], and 30% [30] cut-off level for the Aggregate dimension (AD), Environment (ENV), Social (SOC), and Corporate Governance (CGV). The long-short portfolio is formed by subtracting the high-ranked portfolio returns of a region from the returns on the high-ranked portfolio of another one (H-H). The Pagan and Sossounov (2003) procedure is used in order to identify the different market states. Panel A displays estimates of portfolio financial performance based on equation (3) incorporating a new dummy variable (mixed) that takes value 1 when a market is in a bull period and the other one is in a bear period, and zero otherwise. Global factors are used to estimate the financial portfolio performance across regions. Panel B shows results of controlling for industry effects following the approach of Geczy et al. (2003). 25 TRBC industry global indices are used and principal components with eigenvalues superior to 1 are selected. The multi factor-models are estimated by OLS based on the heteroskedasticity and autocorrelation adjusted errors of Newey and West (1987). The asterisks are used to represent the statistically significant coefficients at the 1% (***), 5% (**) and 10% (*) significance levels. The full sample period is from January 2002 to December 2017.

Panel A: Alphas of the long-short portfolios: H-H

	Both in Bear				Both in Bull				Mixed			
EU & NA	AD	ENV	SOC	CGV	AD	ENV	SOC	CGV	AD	ENV	SOC	CGV
Long-Short [10]	-0.0049	0.0001	-0.0041	0.0075	-0.0034	-0.001	-0.0008	-0.0026	-0.0087	-0.0055	-0.0027	-0.0153**
Long-Short [20]	-0.0065	-0.0069	-0.0078	0.0036	0.0006	-0.0007	0.0011	-0.0011	-0.0058	-0.0045	-0.0002	-0.0116*
Long-Short [30]	-0.0073	-0.0056	-0.0135*	0.0033	0.0002	0.0000	0.0003	-0.0011	-0.0045	-0.0033	-0.0003	-0.0127*
AP & NA	AD	ENV	SOC	CGV	AD	ENV	SOC	CGV	AD	ENV	SOC	CGV
Long-Short [10]	0.0145	0.0142	0.0077	0.0235*	-0.0014	0.0032	-0.0023	0.0018	0.0056	0.0025	0.0086	0.0044
Long-Short [20]	0.0187	0.0105	0.0143	0.0211*	0.0016	0.0008	-0.0006	0.0013	0.0043	0.0034	0.0043	0.0038
Long-Short [30]	0.0131	0.0150	0.0110	0.0233**	0.0001	0.0002	-0.0014	-0.0006	0.0037	0.0005	0.0035	0.0039
JP & NA	AD	ENV	SOC	CGV	AD	ENV	SOC	CGV	AD	ENV	SOC	CGV
Long-Short [10]	0.0018	0.0095	0.0025	0.0050	-0.0044	-0.0039	-0.0022	-0.0048	0.0253***	0.0196***	0.0384***	0.0151***
Long-Short [20]	0.0058	0.0077	0.0015	0.0053	-0.0021	-0.0035	-0.0010	-0.0025	0.0158***	0.0107	0.0215***	0.0418***
Long-Short [30]	0.0007	0.0089	-0.0051	0.0062	-0.0015	-0.0023	-0.0015	-0.0033	0.0262***	0.0325***	0.0374***	0.0408***
AP & EU	AD	ENV	SOC	CGV	AD	ENV	SOC	CGV	AD	ENV	SOC	CGV
Long-Short [10]	0.0045	-0.0044	0.0053	0.0070	0.0022	0.0049	-0.0021	0.0061	0.0472***	0.0431***	0.0460***	0.0491***
Long-Short [20]	0.0051	0.0008	0.0007	0.0041	0.0015	0.0016	-0.0017	0.0032	0.0450***	0.0336***	0.0435***	0.0481***
Long-Short [30]	0.0029	-0.0004	0.0009	0.0065	0.0001	0.0005	-0.0023	0.0009	0.0441***	0.0388***	0.0425***	0.0483***
EU & JP	AD	ENV	SOC	CGV	AD	ENV	SOC	CGV	AD	ENV	SOC	CGV
Long-Short [10]	-0.0067	-0.0093	-0.0066	0.0025	0.0044	0.0045	0.0064	0.0041	-0.0084	-0.0032	-0.0096	-0.0115
Long-Short [20]	-0.0123	-0.0146	-0.0093	-0.0017	0.0066	0.0051	0.0073	0.0038	-0.0072	-0.0048	-0.0058	-0.0110
Long-Short [30]	-0.0080	-0.0145	-0.0084	-0.0030	0.0053	0.0049	0.0064	0.0047	-0.0063	-0.0038	-0.0053	-0.0102
AP & JP	AD	ENV	SOC	CGV	AD	ENV	SOC	CGV	AD	ENV	SOC	CGV
Long-Short [10]	0.0127	0.0048	0.0052	0.0184*	0.0033	0.0063	0.0019	0.0068	0.0144**	0.0076	0.0111	0.0084
Long-Short [20]	0.0129	0.0028	0.0128	0.0158	0.0042	0.0036	0.0019	0.0036	0.0094	0.0069	0.0096	0.0062
Long-Short [30]	0.0124	0.0061	0.0161	0.0171*	0.0021	0.0019	0.0013	0.0024	0.0076	0.0069	0.0063	0.0094*

Panel B: Alphas of the long-short portfolios controlling for industry effects: H-H

	Both in Bear				Both in Bull				Mixed			
EU & NA	AD	ENV	SOC	CGV	AD	ENV	SOC	CGV	AD	ENV	SOC	CGV
Long-Short [10]	0.0099	0.0175*	0.0111	0.026***	-0.0040	-0.0019	-0.0011	-0.0038	-0.0136*	-0.0106	-0.0086	-0.0190***
Long-Short [20]	0.0119	0.0067	0.0056	0.0203**	0.0001	-0.0013	0.0012	-0.0022	-0.0120*	-0.0094	-0.0068	-0.0153**
Long-Short [30]	0.0086	0.0077	-0.0005	0.0204**	-0.0002	-0.0005	0.0001	-0.0019	-0.0102	-0.0080	-0.006	-0.0173**
AP & NA	AD	ENV	SOC	CGV	AD	ENV	SOC	CGV	AD	ENV	SOC	CGV
Long-Short [10]	0.0057	0.0156	0.0035	0.0292*	-0.0015	0.0037	-0.0020	0.0015	0.0107	-0.0018	0.0077	0.0033
Long-Short [20]	0.0183	0.0097	0.0104	0.0178	0.0017	0.0011	-0.0005	0.0011	0.0036	0.0016	0.0048	0.0065
Long-Short [30]	0.0080	0.0080	0.0045	0.0193	0.0002	0.0004	-0.0012	-0.0007	0.0052	0.0016	0.0048	0.0060
JP & NA	AD	ENV	SOC	CGV	AD	ENV	SOC	CGV	AD	ENV	SOC	CGV
Long-Short [10]	0.0131	0.0209	0.0175	0.022	-0.0059	-0.0053	-0.0043	-0.0071**	0.0261**	0.0174	0.0388***	0.0261***
Long-Short [20]	0.0257*	0.0237	0.0230	0.0206	-0.0048	-0.0056	-0.0039	-0.0046	0.0233**	0.0144	0.0266**	0.0472***
Long-Short [30]	0.0198	0.0267*	0.0150	0.0217*	-0.0041	-0.0047	-0.0042	-0.0053	0.0325***	0.0371***	0.0431***	0.0455***
AP & EU	AD	ENV	SOC	CGV	AD	ENV	SOC	CGV	AD	ENV	SOC	CGV
Long-Short [10]	0.0084	-0.0030	0.0072	0.0102	0.0014	0.0054	-0.0027	0.0057	0.0526***	0.0435***	0.0508***	0.0517***
Long-Short [20]	0.0082	0.0034	0.0046	0.0070	0.0009	0.0012	-0.0030	0.0029	0.0498***	0.0377***	0.0520***	0.0517***
Long-Short [30]	0.0055	0.0019	0.0040	0.0090	-0.0006	-0.0002	-0.0032	0.0003	0.0497***	0.0450***	0.0491***	0.0531***
EU & JP	AD	ENV	SOC	CGV	AD	ENV	SOC	CGV	AD	ENV	SOC	CGV
Long-Short [10]	-0.0086	-0.0116	-0.0085	-0.0030	0.0046	0.0051	0.0067	0.0045	-0.0100	-0.0054	-0.0107	-0.0130*
Long-Short [20]	-0.0185	-0.0221	-0.0213	-0.0078	0.0073	0.0059	0.0087	0.0038	-0.0086	-0.0061	-0.0071	-0.0112
Long-Short [30]	-0.0163	-0.0234*	-0.0196	-0.0092	0.0061	0.0058	0.0076	0.0049	-0.0074	-0.0050	-0.0065	-0.0105
AP & JP	AD	ENV	SOC	CGV	AD	ENV	SOC	CGV	AD	ENV	SOC	CGV
Long-Short [10]	-0.0044	-0.0085	-0.0068	0.0086	0.0038	0.0071	0.0027	0.0077	0.0196**	0.0099	0.0131	0.0089
Long-Short [20]	-0.0034	-0.0124	-0.0079	-0.0004	0.0050	0.0043	0.0028	0.0042	0.0130*	0.0103	0.0142**	0.0106
Long-Short [30]	-0.0080	-0.0150	-0.0056	-0.0007	0.0031	0.0032	0.0024	0.0031	0.0119*	0.0104*	0.0105	0.0138**

4.5 Discussion and conclusions

In this paper we investigate the financial performance of international stock portfolios based on CSR criteria. Using an international dataset of companies between 2002 and 2017, we extend the evidence on SRI portfolio performance to North America, Europe, Japan, and Asia Pacific. An initial look at the data discloses that number of firms that are rated according to their ESG concerns across regions has increased progressively, suggesting an increase in the market-investor demand for this kind of information. It represents an additional evidence of the growing interest of investors for knowing extra-financial information (ESG) of firms.

We consider different screening effects by evaluating portfolios formed both on an aggregate dimension of CSR and on specific ESG dimensions. We document important differences in the ESG scores of firms across regions. NA firms are particularly concerned on Governance issues; EU firms are paying further attention to Social demands; both EU and JP firms are more sensitive to Environmental aspects; and JP firms do not perform so well on Governance issues. This evidence is in line with Louche and Lydenberg (2006), and Sakuma and Louche (2008), who document regional and cultural idiosyncrasies in socially responsible investing.

Using a multi-factor model controlling for industry effects, we first estimate alphas of long-short portfolios under different cut-offs within regions. Our results show that, in most cases, there are no statistical significant differences in the performance of portfolios of firms with higher sustainability scores and those with lower sustainability scores. Nevertheless, we find particular influential effects of screening processes within regions. In terms of the Governance dimension, we observe contrasting effects: some EU high-rated portfolios underperform low-rated ones, while some AP high-rated portfolios outperform low-rated ones. On the Social dimension, we find a strong positive effect in EU firms: high-ranked firms outperform their low-ranked counterparts whatever the cut-off used. Screening processes based on the Environment dimension do not uncover significant performance differences. This evidence is in line with with Eccles et al. (2011), and Cortez et al. (2012), who identify that the concerns and investment styles of investors are different across regions, and also with Louche and Lydenberg (2006), and Neher and Hebb (2015) who suggest that regional and cultural differences may affect financial performance of socially responsible investing. We also compare the financial performance of high-rated SRI portfolios among regions. The results show that, in general, high-ranked firms of different regions perform similarly.

This suggests that regional-specific aspects do not seem to have a significant effect on the financial performance of high-ranked SRI firms when they are compared to each other.

In addition, we analyse how SRI portfolios perform across different market states. We estimate the alphas of regional portfolios over different market phases and find mixed results. While for the full sample period we found just a few cases with differences in performance between high- and low-rated portfolios, assessing the financial performance over different market phases uncovers several portfolio performance differentials. The differences are especially notable in EU over bull markets. Several portfolios of high-ranked firms on the Aggregate, Environment, and especially the Social dimension outperform their low-ranked peers. In the AP market, controlling for industry effects uncovers a significant outperformance of well-governed firms during bear markets. Additionally, we find an adverse impact of social screens in bear markets on the Environment dimension in NA firms after adjustment for influential industry characteristics. Since bull and bear periods do not always match across markets, we further establish a third market phase to evaluate portfolio performance differences among SRI firms in times where different regions are experiencing different market cycles, Our findings suggest that significant performance differences among SRI portfolios of different regions is a result of firms of different regions being influenced by distinct economic conditions. In sum, these results suggest that the impact of social screening in portfolio performance is market state and geographically dependent and are in line with Badía et al. (2017), who also document that the financial performance of SRI portfolios is market state dependent.

In general, our results reflect the different patterns of ESG interests across different regions. Clearly, the Social dimension is relevant to EU SRI investors. EU firms score highest on this dimension and high-rated portfolios formed on this dimension perform better than low-rated ones. This result suggests that European markets are rewarding socially responsible investors that screen on this dimension. The Governance dimension is relevant both for EU firms and for AP firms. For the former, the effect of screening on this dimension is negative and for the latter it is positive. This suggests that investors in these regions are in different stages of maturity regarding their awareness and understanding of the effects of governance issues in the firm valuation.

Our evidence supports the view that the consideration of specific dimensions of CSR is useful (e.g., Van de Velde et al., 2005; Galema et al., 2008; Hoepner et al., 2016).

Different ESG screens have differential impacts on financial performance of portfolios across regions and our results suggest that looking at different dimensions of CSR is useful for investors who wish to 'do good while doing well'. Our results also highlight how industry influences affect the relationship between CSR and financial performance. We observe that industry exposure affects portfolios in some regions and especially under different market conditions. These findings are consistent with those of Porter and Kramer (2006) and Hoepner et al. (2010), who find that industry characteristics even drive the financial performance of SRI portfolios.

Overall, our findings support the argument that, in general, global investors can align their personal concerns related to social and ethical values and beliefs with their investment decisions without sacrificing financial performance. They can even benefit from abnormal returns if investing in companies of specific geographies according to specific dimensions of social responsibility and under different market conditions. Besides making the case for investing with a conscience, our results also suggest that SRI can be used as an investment process to change and improve the behaviour of corporations in different regions. Professional investment managers driving their funds towards responsible firms yield similar or higher financial performance compared to a conventional investment approach and, in addition, they can attract an increasing segment of investors concerned with SRI demands.

References

Asem, E. (2009). Dividends and price momentum. Journal of Banking & Finance, 33(3), 486-494.

Auer, B. R. (2016). Do socially responsible investment policies add or destroy European stock portfolio value? Journal of Business Ethics, 135(2), 381-397.

Auer, B. R., & Schuhmacher, F. (2016). Do socially (ir) responsible investments pay? new evidence from international ESG data. The Quarterly Review of Economics and Finance, 59, 51-62.

Badía, G., Ferruz, L., and Cortez, M. (2017, September). SRI to retail investors in international stock markets and on different market states. XIX AECA International Conference, Santiago de Compostela, Spain.

Becchetti, L., Ciciretti, R., Dalò, A., & Herzel, S. (2015). Socially responsible and conventional investment funds: Performance comparison and the global financial crisis. Applied Economics, 47(25), 2541-2562.

Bollen, N. P. (2007). Mutual fund attributes and investor behavior. Journal of Financial and Quantitative Analysis, 42(3), 683-708.

Borgers, A., Derwall, J., Koedijk, K., & Ter Horst, J. (2013). Stakeholder relations and stock returns: On errors in investors' expectations and learning. Journal of Empirical Finance, 22, 159-175.

Boutin-Dufresne, F., & Savaria, P. (2004). Corporate social responsibility and financial risk. The Journal of Investing, 13(1), 57-66.

Brammer, S., Brooks, C., & Pavelin, S. (2006). Corporate social performance and stock returns: UK evidence from disaggregate measures. Financial Management, 35(3), 97-116.

Brammer, S., Brooks, C., & Pavelin, S. (2009). The stock performance of America's 100 best corporate citizens. The Quarterly Review of Economics and Finance, 49(3), 1065-1080.

Brzeszczynski, J., & McIntosh, G. (2014). Performance of portfolios composed of British SRI stocks. Journal of Business Ethics, 120(3), 335.

Capelle-Blancard, G., & Monjon, S. (2012). Trends in the literature on socially responsible investment: Looking for the keys under the lamppost. Business Ethics: A European Review, 21(3), 239-250.

Carhart, M. (1997). On persistence in mutual fund performance. The Journal of Finance, 52(1), 57-82.

Carvalho, A., & Areal, N. (2016). Great places to work®: Resilience in times of crisis. Human Resource Management, 55(3), 479-498.

Cooper, M. J., Gutierrez Jr, R. C., & Hameed, A. (2004). Market states and momentum. The Journal of Finance, 59(3), 1345-1365.

Cortez, M. C., Silva, F., & Areal, N. (2009). The performance of European socially responsible funds. Journal of Business Ethics, 87(4), 573-588.

Cortez, M. C., Silva, F., & Areal, N. (2012). Socially responsible investing in the global market: The performance of US and European funds. International Journal of Finance & Economics, 17(3), 254-271.

Derwall, J., Guenster, N., Bauer, R., & Koedijk, K. (2005). The Eco-Efficiency Premium Puzzle. Financial Analysts Journal, 61(2), 51-63.

Derwall, J., Koedijk, K., & Ter Horst, J. (2011). A tale of values-driven and profit-seeking social investors. Journal of Banking & Finance, 35(8), 2137-2147.

DiBartolomeo, D., & Kurtz, L. (1999). Managing risk exposures of socially screened portfolios. Northfield Information Services, 1-17.

Eccles, R. G., Ioannou, I., & Serafeim, G. (2014). The impact of corporate sustainability on organizational processes and performance. Management Science, 60(11), 2835-2857.

Eccles, R. G., Serafeim, G., & Krzus, M. P. (2011). Market interest in nonfinancial information. Journal of Applied Corporate Finance, 23(4), 113-127.

Edmans, A. (2011). Does the stock market fully value intangibles? employee satisfaction and equity prices. Journal of Financial Economics, 101(3), 621-640.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. Journal of Financial Economics, 33(1), 3-56.

Fama, E. F., & French, K. R. (2012). Size, value, and momentum in international stock returns. Journal of Financial Economics, 105(3), 457-472.

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. Journal of Financial Economics, 116, 1-22.

Fama, E. F., & French, K. R. (2017). International tests of a five-factor asset pricing model. Journal of Financial Economics, 123(3), 441-463.

Fama, E. F., & French, K. R. (2018). Comparing Cross-Section and Time-Series Factor Models . Chicago Booth Research Paper No. 18-08; Fama-Miller Working Paper. Available at SSRN: https://ssrn.com/abstract=3255748 or http://dx.doi.org/10.2139/ssrn.3255748

Filbeck, G., & Preece, D. (2003). Fortune's best 100 companies to work for in america: Do they work for shareholders? Journal of Business Finance & Accounting, 30(5-6), 771-797.

Filbeck, G., Gorman, R., & Zhao, X. (2009). The "Best corporate citizens": Are they good for their shareholders? Financial Review, 44(2), 239-262.

Freeman, R. (1984). Strategic Management: A Stakeholder Perspective (Piman, Boston).

Friedman, M. (1970). The social responsibility of business is to increase its profits. New York Times, September 13, 122-126.

Galema, R., Plantinga, A., & Scholtens, B. (2008). The stocks at stake: Return and risk in socially responsible investment. Journal of Banking & Finance, 32(12), 2646-2654.

Geczy, C. C., Stambaugh, R. F., & Levin, D. (2003). Investing in Socially Responsible Mutual Funds. Working paper, University of Pennsylvania. Available at SSRN: abstract 416380. Accessed in January 2018.

Global Sustainable Investment Alliance – GSIA. (2014). Global Sustainable Investment Review 2014. Accessed in January 2018, available at www.gsi-alliance.org

Global Sustainable Investment Alliance – GSIA. (2016). Global Sustainable Investment Review 2016. Accessed in January 2018, available at www.gsi-alliance.org

Halbritter, G., & Dorfleitner, G. (2015). The wages of social responsibility—where are they? A critical review of ESG investing. Review of Financial Economics, 26, 25-35.

Henke, H. M. (2016). The effect of social screening on bond mutual fund performance. Journal of Banking & Finance, 67, 69-84.

Hoepner, A. G., Yu, P. S., & Ferguson, J. (2010). Corporate Social Responsibility across Industries: When can who do well by doing good? Working paper, University of St. Andrews. Available at SSRN: abstract 1284703. Accessed in January 2018

Hoepner, A., Oikonomou, I., Scholtens, B., & Schröder, M. (2016). The effects of corporate and country sustainability characteristics on the cost of debt: An international investigation. Journal of Business Finance & Accounting, 43(1-2), 158-190.

Hong, H., & Kacperczyk, M. (2009). The price of sin: The effects of social norms on markets. Journal of Financial Economics, 93, 15-36.
Hörisch, J., Ortas, E., Schaltegger, S., & Álvarez, I. (2015). Environmental effects of sustainability management tools: An empirical analysis of large companies. Ecological Economics, 120, 241-249.
Humphrey, J. E., Lee, D. D., & Shen, Y. (2012). Does it cost to be sustainable? Journal of Corporate Finance, 18(3), 626-639.
JSIF - Japan Sustainable Investment Forum (2017). White Paper on Sustainable Investment in Japan 2017. Accessed in November 2018, available at www.jsif.jp.net
Javed, M., Rashid, M. A., & Hussain, G. (2016). When does it pay to be good–A contingency perspective on corporate social and financial performance: Would it work? Journal of Cleaner Production, 133, 1062-1073.
Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. The Journal of Finance, 48(1), 65-91.
Jensen, M. (2001). Value maximisation, stakeholder theory, and the corporate objective function. European Financial Management, 7(3), 297-317.
Kempf, A., & Osthoff, P. (2007). The effect of socially responsible investing on portfolio performance. European Financial Management, 13(5), 908-922.
Lahiri, S. (2003). Resampling methods for dependent data, Springer, New York.
Lee, J., Yen, P., & Chan, K. C. (2013). Market states and disposition effect: Evidence from Taiwan mutual fund investors. Applied Economics, 45(10), 1331-1342.
Leite, P., & Cortez, M. C. (2015). Performance of European socially responsible funds during market crises: Evidence from France. International Review of Financial Analysis, 40, 132-141.
Louche, C., & Lydenberg, S. (2006, July). Socially responsible investment: Differences between Europe and the United States. In Proceedings of the International Association for Business and Society (Vol. 17, pp. 112-117).
Lu, W., Chau, K., Wang, H., & Pan, W. (2014). A decade's debate on the nexus between corporate social and corporate financial performance: A critical review of empirical studies 2002–2011. Journal of Cleaner Production, 79, 195-206.
Margolis, J. D., & Walsh, J. P. (2003). Misery loves companies: Rethinking social initiatives by business. Administrative Science Quarterly, 48(2), 268-305.
Margolis, J. D., Elfenbein, H. A., & Walsh, J. P. (2009). Does it pay to be good... and does it matter? A meta-analysis of the relationship between corporate social and financial performance. Working paper. Available at SSRN: abstract 1866371. Accessed in January 2018.
Mollet, J. C., & Ziegler, A. (2014). Socially responsible investing and stock performance: New empirical evidence for the US and European stock markets. Review of Financial Economics, 23(4), 208-216.
Mollet, J. C., von Arx, U., & Ilić, D. (2013). Strategic sustainability and financial performance: exploring abnormal returns. Journal of Business Economics, 83(6), 577-604.
Neher, A. L., & Hebb, T. (2015). The responsible investment atlas–an introduction. The Routledge Handbook of Responsible Investment, 53-57. Routledge, New York.
Nofsinger, J., & Varma, A. (2014). Socially responsible funds and market crises. Journal of Banking & Finance, 48, 180-193.
Orlitzky, M., Schmidt, F. L., & Rynes, S. L. (2003). Corporate social and financial performance: A meta-analysis. Organization Studies, 24(3), 403-441.
Pagan, A. R., & Sossounov, K. A. (2003). A simple framework for analysing bull and bear markets. Journal of Applied Econometrics, 18(1), 23-46.

Porter, M. E., & Kramer, M. R. (2006). The link between competitive advantage and corporate social responsibility. Harvard Business Review, 84(12), 78-92.
Preston, L. E., & O'bannon, D. P. (1997). The corporate social-financial performance relationship: A typology and analysis. Business & Society, 36(4), 419-429.
Revelli, C., & Viviani, J. (2015). Financial performance of socially responsible investing (SRI): What have we learned? A meta-analysis. Business Ethics: A European Review, 24(2), 158-185.
Sakuma, K., & Louche, C. (2008). Socially responsible investment in Japan: Its mechanism and drivers. Journal of Business Ethics, 82(2), 425-448.
Statman, M., & Glushkov, D. (2009). The wages of social responsibility. Financial Analysts Journal, 65(4), 33-46.
Statman, M., & Glushkov, D. (2016). Classifying and measuring the performance of socially responsible mutual funds. The Journal of Portfolio Management, 42(2), 140-151.
Utz, S., & Wimmer, M. (2014). Are they any good at all? A financial and ethical analysis of socially responsible mutual funds. Journal of Asset Management, 15(1), 72-82.
Van de Velde, E., Vermeir, W., & Corten, F. (2005). Corporate social responsibility and financial performance. Corporate Governance: The international journal of business in society, 5(3), 129-138.
Wimmer, M. (2013). ESG-persistence in socially responsible mutual funds. Journal of Management and Sustainability, 3(1), 9-15.
Woodward, G., & Anderson, H. M. (2009). Does beta react to market conditions? Estimates of ‘bull’ and ‘bear’ betas using a nonlinear market model with an endogenous threshold parameter. Quantitative Finance, 9(8), 913-924.

Chapter 5: Financial performance of government bond portfolios based on environmental, social and governance criteria

Guillermo Badía, Vicente Pina, and Lourdes Torres

Faculty of Economics and Business, University of Zaragoza,
Zaragoza 50005, Spain.

Abstract

We evaluated the financial performance of government bond portfolios formed according to socially responsible investment (SRI) criteria. We thus open a discussion on the financial performance of SRI for government bonds. Our sample includes 24 countries over the period of June 2006 to December 2017. Using various financial performance measures, the results suggest that high-rated government bonds, according to environmental, social, and governance (ESG) dimensions, outperform low-ranked bonds under any cut-off, although differences are not statistically significant. These findings suggest that ESG screenings can be used for government bonds without sacrificing financial performance.

Keywords Socially responsible investments; Government bonds; International finance; Performance evaluation

This study is published in the journal *Sustainability*, journal indexed in the Journal of Citation Report (JCR): Badía, G., Pina, V., & Torres, L. (2019). Financial performance of government bond portfolios based on environmental, social and governance criteria. *Sustainability*, 11(9), 2514.

5.1 Introduction

The growth in socially responsible investment (SRI) has been notable. According to the 2016 Global Sustainable Investment Review, in 2016, US$22.89 trillion of assets were being professionally managed under responsible investment strategies worldwide, an increase of 25% since 2014. In 2016, 53% of managers in Europe used responsible investment strategies, this proportion being 22% in the U.S. and 51% in Australia and New Zealand. Perthe Global Sustainable Investment Association (GSIA) 2016, sustainable investing is an investment approach that considers environmental, social, and governance (ESG) factors in portfolio selection and management. ESG screening investment processes, which allow an investor to select or exclude investments from the available universe based on ESG criteria, have helped investors to align their personal beliefs and values with their investment decisions. Rising individual awareness of environmental, social, and ethical concerns is now strongly influencing the purchasing decisions of investors (Mollet and Ziegler, 2014).

The concept of SRI was originally related to stock selection. However, the proportion of portfolio investors applying SRI criteria to bonds has grown significantly. According to the European Sustainable Investment Forum (EUROSIF, 2016), equities represented over 30% of SRI assets in December 2015, a significant decrease from the previous year's 50%. A strong increase in bonds simultaneously occurred from the 40% registered in December 2013 to 64% in December 2015. Both corporate bonds and government bonds underwent remarkable growth. The former rose from 21.3% to 51.17% of the bond allocation, while the latter increased from 16.6% to 41.26%.

In this regard, the financial implications of the ESG screening processes on corporate bonds may be closely related to stock selections, since corporate bonds are associated with firms. Previous studies (Derwall and Koedijk, 2009; Leite and Cortez, 2016), which evaluated the financial performance of mutual funds that invested in socially responsible fixed-income stocks, found that the average SRI bond funds performed similarly to conventional funds. These results are in line with most empirical studies about the performance of SRI funds, which showed that they tend to perform similarly to their conventional peers (Revelli and Viviani, 2015). However, the ESG screening processes for government bonds, since they are not related to firms, can help provide an in-depth understanding of SRI consequences for alternative assets. Despite the SRI government bond market growth and the development of country ratings based on ESG

factors, the link between government bond returns and country performance in terms of ESG concerns has been overlooked. To the best of our knowledge, no previous research has evaluated the financial performance of responsible government bond investments.

The main objective of this paper was to fill this gap. We evaluated the financial performance of government bond portfolios formed according to ESG criteria. In contrast withprevious studies, which applied firm sustainability ratings, we used sustainability ratings related to countries. We employed the RobecoSAM country sustainability ranking developed by RobecoSAM and Robeco. This ranking is a comprehensive framework for assessing countries' ESG performance. By focusing on ESG factors, such as aging, competitiveness, and environmental risks, this country sustainability ranking offers a view of a country's strengths and weaknesses.

Previous research has shown that ESG factors are valuable for government bonds. Capelle-Blancard et al. (2016) assessed whether ESG performance influences government bond spreads. They found that countries with good ESG performance tended to have less default risk and thus lower bond spreads. Hence, the findings of Hoepner and Neher (2016) were reinforced. They found a negative and significant relationship between government debt and a national sustainability rating. We wanted to ascertain whether ESG factors are valuable from a portfolio management perspective. Drut (2010) assessed a feasible diversification portfolio problem associated with government bond portfolios. They computed the efficient frontier of portfolios, including government bonds from 20 developed countries, and showed that government bond portfolios with high social responsibility scores could be formed without significant loss of diversification. Investors could thus form government bond portfolios based on socially responsible ratings without renouncing the potential for diversification. We wanted to complete a deeper examination and determine whether government bond portfolios formed according to ESG dimensions can be formed without sacrificing financial performance.

We therefore contribute to the existing literature on the financial performance of SRI by examining the impact of ESG screening processes on portfolios of government bonds. Ullmann (1985) noted that stakeholders (e.g., investors, customers, and community) have the power to influence management's corporate social responsibility (CSR) activities and strategies. SRI demands have led firms to pay more attention to their CSR activities and strategies. Hence, our study may lead governments to be more concerned about social, governance, and environmental policies. Given the growth of SRI in

international capital markets and the increasing interest of investors in government bonds, our results about the implications of sustainability screening processes on government bonds in an international context are of practical interest for particular and institutional investors, as well as governments worldwide.

The rest of the paper is organized as follows: Section 2 presents a brief literature review on the financial outcomes of SRI for alternative assets. Section 3 describes the data. Section 4 presents and discusses the empirical analysis, and Section 5 summarizes our main findings and presents our concluding remarks.

5.2 Literature review

The growth in SRI and its consequences have stimulated empirical studies assessing financial behaviors. Prior studies mainly evaluated the financial performance of SRI investment funds and SRI stock portfolios. As Osthoff (2015) noted, many studies compared the performance of SRI investment funds with conventional investments (e.g., Kreander et al. 2005; Gregory and Whittaker, 2007; Utz and Wimmer, 2014). In general, these studies found no significant differences between the financial performance of SRI investment funds and conventional funds (Revelli and Viviani, 2015). Goldreyer and Diltz (1999) evaluated the financial performance of U.S. SRI fixed-income funds, invested in both corporate and government bonds. They found that SRI fixed-income funds underperformed their conventional peers. By contrast, 20 years later, Derwall and Koedijk (2009) found that U.S. SRI fixed-income funds performed similarly to conventional funds. In European markets, Leite and Cortez (2016) showed that financial performance was geographically dependent: UK SRI fixed-income funds underperformed conventional funds, German SRI fixed-income funds outperformed conventional ones, and French SRI fixed-income funds showed similar performance to their conventional peers.

Despite all this attention being valuable from a practical point of view, certain limitations are related to fund studies. Brammer et al. (2006), and Kempf and Osthoff (2007) pointed out that confusing effects, such as fund manager performance and management fees, complicate showing differences in investment fund performance. Evidence provided by Utz and Wimmer (2014), Humphrey et al. (2016), and Statman and Glushkov (2016) showed that the ‘socially responsible’ label may be more akin to a marketing strategy, thus raising doubts among investors whether an SRI fund is really socially responsible. As a consequence, investors may struggle to know the extent to

which an SRI fund is really considering social criteria in its selection process. To address these concerns, some studies followed a portfolio stock approach. They formed portfolios, including high- and low-ranked firms according to their ESG scores, and investigated their financial differences. These studies found ambiguous results. Van de Velde et al. (2005), Galema et al. (2008), and Mollet and Ziegler (2014) did not find significant financial differences between high- and low-ranked sustainable firms. Derwall et al. (2005), Kempf and Osthoff (2007), and Eccles et al. (2014) showed that high-rated portfolios outperformed low-rated ones, but Brammer et al. (2006), and Auer and Schuhmacher (2016) found that high-ranked firms underperformed compared to their low-rated counterparts. In this paper, we follow this approach to elude drawbacks related to fund studies.

The financial implications of SRI strategies have led to intensive research on several assets. Surprisingly and despite the growth in investors applying SRI criteria to government bonds, financial performance of SRI government bond portfolios has been overlooked.

5.3 Data

The samples evaluated mainly focused on developed countries, where valuable information exists on government bond returns and concerns related to SRI behavior, such as the country's institutional framework, regulatory quality, rule of law, government efficiency, political stability, social cohesion, orderly conflict resolution, environmental vulnerabilities and policies, energy dependency, etc. Our dataset included 24 countries over the period June 2006 to December 2017. The countries were Australia, Austria, Belgium, Canada, China, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Japan, the Netherlands, New Zealand, Norway, Poland, Portugal, Spain, Sweden, Switzerland, Turkey, the United Kingdom, and the United States. The countries assessed further represent a significant share of the world income economy and international bond markets. According to the World Bank country classifications by income level (2018–2019), the countries assessed belong to the leading group, except China and Turkey, which belong to the upper–middle-income economies. The data on government bond monthly total returns were sourced from FTSE Global Government Bond Indices 'All maturities', downloaded from the Thomson Reuters database in U.S. dollars. To classify government bonds according to ESG performance, we used the RobecoSAM country sustainability ranking. Robeco and RobecoSAM have jointly

developed a comprehensive and systematic framework for determining country sustainability rankings. Sources used by RobecoSAM include international organizations, such as the World Bank, the United Nations, and the International Labor Organization, as well as a variety of reputable government agencies, private institutions, and non-governmental organizations (NGOs). The framework forms the basis for incorporating environmental, social, and governance risk analysis into the construction process for Robeco and RobecoSAM's government debt portfolios and indices. RobecoSAM's country sustainability framework is used to evaluate many countries on the basis of a broad range of ESG factors that are considered key risk and return drivers for investors (see Appendix A for an extensive explanation of sustainability dimensions). It consists of 17 indicators, each of which is based on various data series, or sub-indicators, whereby each indicator is assigned a predefined weight out of the total framework. Based on the standardized scores, and for each of the indicators and their corresponding weights, countries receive a sustainability score ranging from 1 to 10, with 10 being the highest. The resulting scores offer insights into the investment risks and opportunities associated with each country, allowing investors to better compare countries. The weighting scheme is reviewed periodically, reflecting RobecoSAM's view on the potential impact of each indicator on a country's risk profile.

Figures 1 and 2 show the top five and bottom five countries according to the country sustainability ranking for the first (2006, first semester) and last (2017, second semester) periods, respectively. The countries at both the top and the bottom have remained the same despite more than 10years passing between the two classifications. This evidence suggests that a noteworthy traditional and cultural component may exist behind ESG concerns. In this regard, some studies identified that country-specific factors tend to affect the financial performance of SRI (Eccles et al. 2011; Cortez et al. 2012; Hörisch et al. 2015). These figures may highlight a limited capacity of previous policy initiatives to improve ESG standards in low-rated countries.

Figure 5-1. Country sustainability ranking

Top five and bottom five countries in the first semester of 2006. Based on the standardized scores, and for each of the indicators and their corresponding weights, countries receive a sustainability score ranging from 1 to 10, with 10 being the highest (x-axis).

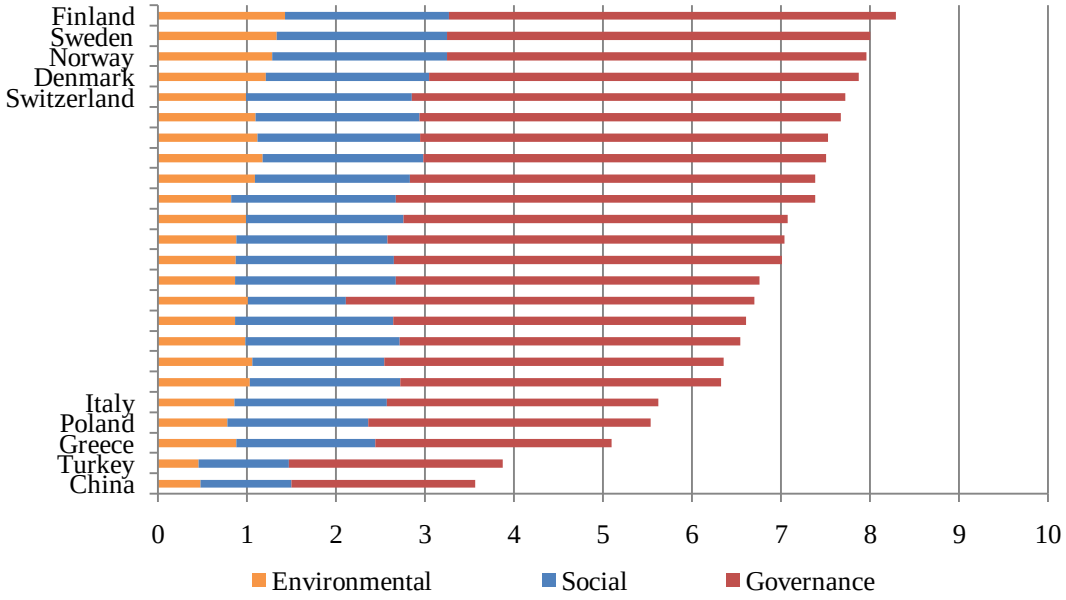
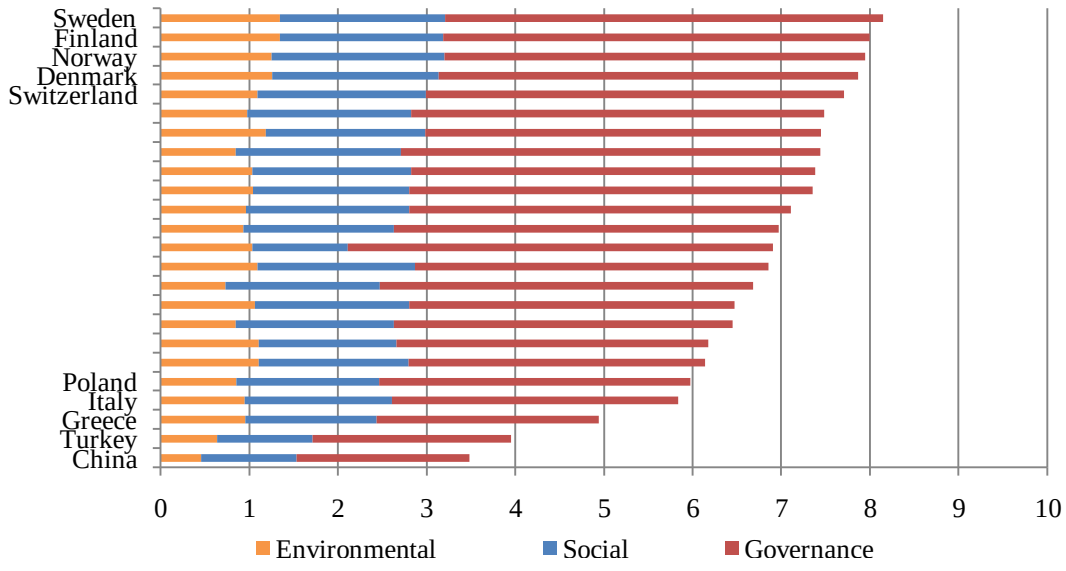


Figure 5-2. Country sustainability ranking

Top five and bottom five countries in the second semester of 2017. Based on the standardized scores, and for each of the indicators and their corresponding weights, countries receive a sustainability score ranging from 1 to 10, with 10 being the highest (x-axis).



5.4 Empirical analysis

5.4.1 Portfolio construction

In this paper, we evaluated the financial implications of social responsibility screenings on government bonds. We ranked government bonds at time *t* according to countries' ESG scores available at *t* –1. We then formed a high- and a low-rated portfolio, including ESG outperformers and underperformers, respectively. Since the country sustainability ranking is updated semi-annually, portfolios are formed twice a year. Related studies applied several cut-offs (Halbritter and Dorfleitner, 2015; Auer 2016). We also used alternative cut-offs (10%, 20%, 30%, 40%, and 50%), which allowed us to evaluate different SRI demand levels. For instance, at the 10% cut-off level, the high-rated portfolio included 10% government bonds from countries with the highest ESG scores, whereas the low-rated portfolio included 10% government bonds from countries with the lowest ESG scores. We formed equally-weighted portfolios rather than value-weighted ones to improve diversification. Drut (2010) showed that highly socially responsible government bond portfolios could be formed without significant loss of diversification. Nonetheless, given the process involved informing a value-weighted portfolio, the standard deviation may have been affected, since this type of portfolio is less diversified. Statman and Glushkov (2009), for instance, found that a value-weighted portfolio (top–bottom portfolio) had a higher standard deviation than an equally-weighted one.

Table 5-1. Descriptive statistics

This table presents a summary statistics of the high- and low-rated portfolios at the 10% (10), 20% (20), 30% (30), 40% (40), and 50% (50) cut-offs.Mean (SD) is the average return (standard deviation) of portfolios. Difference is the mean (SD) difference between high and low portfolios. The full sample period was from June 2006 to December 2017.

	High (10)	Low (10)	High (20)	Low (20)	High (30)	Low (30)	High (40)	Low (40)	High (50)	Low (50)
Mean	0.0026	0.0013	0.0028	0.0026	0.0039	0.0034	0.0039	0.0034	0.0038	0.0035
Difference	0.0013		0.0002		0.0005		0.0005		0.0003	
SD	0.0298	0.0310	0.0282	0.0319	0.0284	0.0324	0.0279	0.0287	0.0266	0.0271
Difference	–0.0012		–0.0038		–0.0040		–0.0008		-0.0005	

Table 1 provides descriptive statistics for the high- and low-rated portfolios at different cut-offs. The high-rated portfolios showed higher average returns than the low-rated ones at any cut-off. As for standard deviation, the high-ranked government bonds showed lower variability in terms of returns. This evidence suggests that risk affected low-rated portfolios to a larger extent than high-rated ones. Descriptive statistics allowed us to identify what the financial outcomes of ESG screening processes on

government bonds may be. However, an extensive evaluation using risk-adjusted measures is advisable.

5.4.2 Ledoit and Wolf approach

To estimate statistical financial performance differences between high and low portfolios, we followed the Ledoit and Wolf (2008) (LW) approach. Accordingly, the Sharpe ratio (1966)—the ratio of excess return to standard deviation—was used to compare the performance of alternative investment strategies. From two investment portfolios, *i* and *j*, whose excess returns over the risk-free rate at time *t* were r_{ti} and r_{tj}, respectively, a total of T return pairs $(r_{1i},\ r_{1j}),\ldots,(r_{Ti},\ r_{Tj})$ were observed. The difference between two Sharpe ratios is given by $\Delta = \mathrm{Sh}_i - \mathrm{Sh}_j = \mu_i/\sigma_i - \mu_j/\sigma_j$, where μ and σ are the sample mean and standard deviation, respectively. To run statistical inference between the two Sharpe values, prior studies (Demiguel and Nogales, 2009; Gasbarro et al. 2007) used the Jobson and Korkie (1981) test and the correction proposed by Memmel (2003). However, this test is not valid if the returns distribution is non-normal, or if the observations are correlated over time, both phenomena being quite common in financial returns time series data. LW proposed a studentized time series bootstrap approach that works asymptotically and has satisfactory properties in finite samples. The literature (Hall, 1992; Lahiri, 2003) shows the enhanced inference accuracy of the studentized bootstrap over standard inference based on asymptotic normality. LW proposed testing $H_0: \Delta = \mathrm{Sh}_i - \mathrm{Sh}_j = 0$ by inverting a bootstrap confidence interval. A two-sided bootstrap confidence interval with nominal level 1–α for Δ (true difference between the Sharpe ratios) was constructed and if zero was not contained in the interval, then $H_0$was rejected at nominal level α. Specifically, LW proposed constructing a symmetric studentized time series bootstrap confidence interval. To do this, the two-sided distribution function of the studentized statistic is approximated through the bootstrap by $\mathrm{F}\left(\left|\hat{\Delta} - \Delta\right|/s(\hat{\Delta})\right) \approx F\left(\left|\hat{\Delta}^* - \Delta\right|/s(\hat{\Delta}^*)\right)$, where Δ is the true difference between the Sharpe ratios, $\hat{\Delta}$is the estimated difference computed from the original data, $s(\hat{\Delta})$ is a standard error for $\hat{\Delta}$(also calculated from the original data), $\hat{\Delta}^*$ is the estimated difference computed from bootstrap data, and $s(\hat{\Delta}^*)$ is a standard error for $\hat{\Delta}^*$ (also calculated from bootstrap data). Letting $z^*_{|\cdot|,\lambda}$ be a λ quantile of $\mathrm{F}\left(\left|\hat{\Delta}^* - \Delta\right|/s(\hat{\Delta}^*)\right)$, a bootstrap 1−α confident

interval for Δ is given by $\hat{\Delta} \pm z^{*}_{|\cdot|,1-\alpha} s(\hat{\Delta})$. LW noted that with heavy-tailed data or data of a time series nature, this quantile will typically be somewhat larger than its standard normal counterpart (used in the traditional tests) in small to moderate samples, resulting in more conservative inferences. To generate the bootstrap data, we used the circular block bootstrap of Politis and Romano (1992), resampling blocks of pairs from the observed pairs (r_{ti}, r_{tj}), t = 1,…, T, with a replacement. Applying the studentized circular block bootstrap requires a choice of the block size b, and LW proposed using the calibration procedure of Loh (1987), suggesting that M = 5000 bootstrap sequences is sufficient for reliable inference. The standard error $s(\hat{\Delta})$ is calculated using kernel estimation, specifically the pre-whitened quadratic spectral kernel of Andrews and Monahan (1992). The standard error $s(\hat{\Delta}^{*})$ is the natural standard error calculated from the bootstrap data, making use of a special block dependence structure. The bootstrap p-values are computed as $PV = \{\tilde{d}^{*,m} \geq d\} + 1/M + 1$, where $d = |\hat{\Delta}|/s(\hat{\Delta})$, the original studentized test statistic, $\tilde{d}^{*,m} = |\hat{\Delta}^{*,m} + \hat{\Delta}|/s(\hat{\Delta}^{*,m})$, denotes the centered studentizedstatistic computed from the mth bootstrap sample by $d^{*,m}$, m = 1,…, M, and M is the number of bootstrap resamples.

Table 2 shows the results of applying the Sharpe ratio and the LW procedure to estimate the statistical significance of the difference between the Sharpe ratio in high- and low-rated portfolios. We found that high-rated portfolios outperformed low-rated ones with any cut-off. Nonetheless, the LW t-statistic indicated that differences were not statistically significant. These results were in line with most previous studies, which reported that SRI performed similarly to conventional investments. Derwall and Koedijk (2009) found that U.S. SRI fixed-income funds performed similarly to conventional funds. Leite and Cortez (2016) found similar results for German and French SRI fixed-income funds. Nonetheless, these studies included the performance of corporate bonds in their investigations. Although significant differences were not found using different cut-offs, we found that the biggest difference between the Sharpe value of the high- and low-portfolios appeared at the most demanding SRI level, the 10% cut-off. This evidence suggested that government bonds from countries with the best ESG practices performed substantially better than those with the worst practices. Hence, investors driving funds to, for example, countries with a stable institutional framework, high regulatory quality, no environmental vulnerabilities, or nonexistent social conflicts, not only reduced the risks associated with investments, but also achieved financial

performance similar to conventional investments. The superior financial performance of the high-rated portfolios could be seen as a reward for recognizing that countries with outstanding ESG policies should do better than less responsible ones.

Table 5-2. Portfolio financial performance

This table shows portfolio performance of the high- and low-rated portfolios at the 10% (10), 20% (20), 30% (30), 40% (40), and 50% (50) cut-offs based on the Sharpe ratio and the Ledoit and Wolf (LW) significant tests. The Sharpe ratio (SH) and the Ledoit and Wolf (LW) procedure were used to identify statistically significant differences between the Sharpe ratios of the portfolios. Difference is the Sharpe ratio difference between the high- and low-rated portfolios and LW t-test is the t-statistics according to the LW process. The asterisks represent the statistically significant coefficients at the 1% (***), 5% (**), and 10% (*) levels. One-month U.S. T-bills were used as a proxy for the risk-free rate. The full sample period was from June 2006 to December 2017.

	High (10)	Low (10)	High (20)	Low (20)	High (30)	Low (30)	High (40)	Low (40)	High (50)	Low (50)
SH	0.0624	0.0174	0.0736	0.0590	0.1115	0.0837	0.1158	0.0949	0.1169	0.1034
Difference	0.0450		0.0146		0.0278		0.0209		0.0134	
LW t-test	0.6395		0.2842		0.5355		0.4743		0.3298	

5.4.3 Robustness checks

To test the robustness of our findings, we considered three additional financial performance evaluation measures. We used the adaptation proposed by Ferruz and Sarto (2004) (FS) regarding the Sharpe ratio used previously by some studies (Scholz, 2007; Luo et al. 2015). FS noted that the Sharpe ratio assumes positive portfolio excess returns. However, they determined that this was not always the case. Consequently, when this happens, the Sharpe ratio can produce anomalous results. In this context, FS proposed a correction to the Sharpe ratio as follows: $FS_{p,t} = \left(R_{p,t}/R_{f,t}\right)/\sigma_{p,t}$, where $R_{p,t}$ is the portfolio p return at time t, $R_{f,t}$ is the risk-free return at time t, and $\sigma_{p,t}$ is the standard deviation of the portfolio p at time t. We also used the Sortino ratio (Sortino and Van Der Meer, 1991; Sortino and Price, 1994) to evaluate performance on the basis of the lower partial moments (LPM). According to the Sortino ratio, risk is measured by the negative deviations of returns in relation to a minimum acceptable return (e.g., zero, the risk-free rate, or the average return). We used a rolling interest rate based on the evolution of the risk-free monthly interest rate. The Sortino specification is $S_{p,t} = R_{p,t} - \varphi/\left(\frac{1}{T}\sum_{t=1}^{T}\max[\varphi - R_{p,t}, 0]^2\right)^{1/2}$, where $R_{p,t}$ is the portfolio p return at time t, and φ is the target return or minimum acceptable return. This measure has been used previously (Auer, 2016; Leggio and Lien, 2003; Meligkotsidou et al. 2009).

We also computed alpha values from a multi-factor model, including some fiscal and economic variables as controls. Previous related literature (Capelle-Blancard et al. 2016;

Hoepner and Neher, 2016) controlled for gross domestic product (GDP) growth rate, inflation, fiscal condition (debt/GDP and Primary Balance (PB)/GDP), current account, liquidity ratio, country openness, and sovereign credit ratings. For a more extensive discussion and understanding of the effects of each variable, see, for instance Capelle-Blancard et al. (2016). To start, we considered these variables and assessed their significance as determinants of international government bond returns (Appendix B). To this end, we estimated a fixed effects panel data model, as a Hausman test was conducted and showed that a fixed effects model was required instead of a random effects model. This approach was often used to address this concern in previous research. As data on control variables are annual, we used a cubic spline interpolation to generate monthly data. Our results showed that, except for GDP growth rate and debt/GDP, variables were significant and thereby had an impact on government bond returns (data available on request). We consequently left these two variables out of the analysis. We also performed the analysis using the eight control variables and the results were unaltered. The next step to evaluate the financial performance by computing alphas was to include control variables in the multi-factor model. Since we handled six variables for 24 countries, we employed principal component analysis (PCA) to determine the main dimensions. The principal components thus represented a vector of variables capturing fiscal and economic conditions. To end, we formed a long–short portfolio, a difference portfolio, which was formed by subtracting the low-rated portfolio returns from the returns on the high-rated portfolio. The resulting alpha was the estimated financial portfolio performance. This approach was commonly used in previous related studies (Humphrey, et al. 2012; Leite et al. 2018). A challenge in the evaluation of financial performance is the need for controlling alternative explanations. On a corporate side, Dang et al. (2018) studied the use of firm size measures in the literature and found that it is a key variable in this area since affects the independent and dependent variables simultaneously. In this regard, country size measures could affect the financial performance evaluation of bond portfolios. We addressed that point including in the multi-factor model several control variables which could be associated with the size of countries, such as GDP growth rate and current account. In addition, since according to the World Bank country classifications by income level (2018–2019), the countries assessed belong to the leading group, except China and Turkey, which belong to the upper–middle-income economies, our samples mainly focused on developed countries with homogeneous characteristics from a wealth point of view,

thereby restricting country size effects on the financial performance of our bond portfolios.

Finally, a common impediment to understanding the true relationship between different aspects of empirical finance is the endogeneity problem; variables are sometimes endogenous and causality relations are complicated (Li, 2016). Examples of endogeneity problem in our scenery include that bonds which expect to outperform would use SRI, or something not captured in credit rating could affect SRI and performance simultaneously. We evaluate the causality relation between ESG scores and bond performance for each country using the Granger–causality test. Scholtens (2008) is an example of a study that applies this test in a CSR context. We find unidirectional causality from ESG scores to bond performance. In addition, following Li (2016), to deal with a possible endogeneity problem, we include the lagged dependent variable in our multi-factor model used to evaluate financial performance. Our findings do not change significantly. Given that we rank government bonds at *t* according to countries' ESG scores available at *t-1*, the ESG scores used are lagged. Capelle-Blancard et al. 2016, who use lagged ESG scores, suggest that lagging ESG scores helps to avoid the endogeneity problems and simultaneity bias that may arise as a result of a contemporaneous bidirectional causality existing between ESG aspects and bond performance.

Table 3 and 4 display the results of applying the additional portfolio financial performance measures. By using the FS ratio, we found that the results were in line with our previous results. The values of high-rated portfolios were higher than those of the low-rated counterparts at any cut-off. The findings using the Sortino specification were also similar. High-rated portfolios outperformed low-rated ones under any cut-off. Results about these measures were limited to a descriptive comment since processes, such as Ledoit and Wolf (2008), to evaluate statistical significance differences were not available. Finally, we found positive alphas in the long–short portfolios, meaning that high-ranked government bonds outperformed low-ranked ones, although alphas were not significant. These robustness checks supported our previous findings.

Table 5-3. Financial performance using the Ferruz and Sarto ratio and the Sortino ratio

This table presents the financial performance of the high- and low-rated portfolios at the 10% (10), 20% (20), 30% (30), 40% (40), and 50% (50) cut-offs using the Ferruz and Sarto (FS) ratio and the Sortino ratio. One-month U.S. T-bills were used as a proxy for the risk-free rate. The full sample period ranged from June 2006 to December 2017.

	High (10)	Low (10)	High (20)	Low (20)	High (30)	Low (30)	High (40)	Low (40)	High (50)	Low (50)
FS ratio	120.44	56.42	137.92	113.21	190.46	147.2	197.11	166.66	200.35	180.58
Sortino Ratio	0.0906	0.0231	0.1086	0.0796	0.1663	0.114	0.1707	0.1322	0.1716	0.1449

Table 5-4. Financial performance using the multi-factor model

This table presents the financial performance of the long-short portfolios at the 10% (10), 20% (20), 30% (30), 40% (40), and 50% (50) cut-offs using the multi-factor model. A long–short portfolio is the portfolio formed by subtracting high-rated portfolio returns from the low-rated portfolio returns. Six fiscal and economic variables were included in the model using a principal component analysis (PCA) process to control for determinants of international government bond returns. A fixed panel data model was used to select control variables. Models were estimated by ordinary least squares (OLS) based on the heteroskedasticity and autocorrelation adjusted errors of Newey and West (1986). The asterisks represent the statistically significant coefficients at the 1% (***), 5% (**), and 10% (*) significance levels. The full sample period ranged from June 2006 to December 2017.

	Long–Short (10)	Long–Short (20)	Long–Short (30)	Long–Short (40)	Long–Short (50)
Alpha	0.0012	0.0001	0.0004	0.0005	0.0003
t-statistic	0.7485	0.0966	0.3549	0.555	0.3578

5.5 Conclusions

The expansion of SRI has led to extensive research on its financial consequences. Previous research has mainly focused on the financial benefit or the cost of ESG screening processes on investments related to corporate firms. Both mutual funds and stock portfolios have been evaluated from an SRI investment approach. However, despite the growing interest of portfolio investors in applying SRI criteria to government bonds, to the best of our knowledge, no previous studies have investigated the financial outcomes of SRI screenings on government bond portfolios.

Our main objective in this study was to evaluate the financial performance of government bond portfolios formed according to ESG criteria. We opened a discussion on financial performance of SRI for an asset other than firms. Using RobecoSAM information to classify the government bonds according to ESG performance, we assessed financial differences between high- and low-ranked government bonds. Using several portfolio financial performance measures, our results showed that high-rated portfolios outperform low-rated ones under any SRI level (cut-off), although differences were not significant. These findings are in line with most previous studies that reported

that SRI performs similarly to conventional investments. Most empirical studies on the performance of SRI mutual funds across different geographical areas found no significant differences between their performance and that of conventional funds (Leite and Cortez, 2016; Leite et al. 2018). Likewise, many empirical studies evaluating differences between high- and low-ranked firms, according to their CSR scores, also found that the differences are not significant (Mollet and Ziegler, 2014; Halbritter and Dorfleitner, 2015). Therefore, the absence of significant differences is considered a relevant finding in most previous research.

Overall, our evidence indicates that an investor can satisfy ESG concerns without sacrificing financial performance by investing in government bonds. In this regard, as SRI investor claims have led firms to be more concerned with their corporate social responsibility strategies (Mollet and Ziegler, 2014; Ullmann, 1985), for instance, Li et al. (2019) found that SRI mutual funds had a positive effect on firm's future CSR, investors screening government bonds according to their sustainability scores could influence countries in terms of ESG guiding principles. Our results suggest that SRI can be used as a tool to enhance the ESG policies of countries. Currently, many countries are shifting toward a sustainable economy. For instance, the Paris Agreement aims to strengthen the global response to the threat of climate change in the context of sustainable development and efforts to eradicate poverty. To this end, it recommends that financial resources flow toward climate-resilient development and the reduction in greenhouse gas emissions. Global capital markets are one of the most powerful tools in the fight against climate change and to develop sustainable economies. However, they are often overlooked by governments. If governments are aware that social responsibility issues may influence investment decisions, and that investors can satisfy their social concerns and simultaneously produce similar financial performance as conventional investments, then they should improve ESG standards and display this information to attract new investments. Since SRI investors drive their funds toward investments with high levels of sustainability (Mollet and Ziegler, 2014), governments could use the ESG information as a tool to attract an increasing number of investors concerned with SRI issues. Aiming to make the country's interests related to socially responsible concerns visible for investors—in line with the European initiative (Directive 2014/95/EU of the European Parliament), whose objective is that large firms disclose both financial and non-financial information—it might be beneficial for governments to publish official reports about their achievements in socially responsible

policies, strategies, and activities to help SRI investors make well-informed investment decisions. Socially responsible policies and strategies may positively affect bond performance through different channels. Environmental challenges are a potential risk for investors, as environmental externalities can result in significant economic losses, while repairing environmental damage such as air and water pollution can generate considerable fiscal costs. Adequate investments towards preventing environmental problems limit such potential liabilities. Likewise, a weak social climate dominated by labour unrest, extreme inequality or other social tensions is another potential investment risk. Social policies providing a strong social cohesion support orderly conflict resolution and facilitate the implementation of necessary reforms, thus contributing to sustainable economic development.

We consider that further research would be worthwhile to broaden the knowledge in this field, for instance, to evaluate the particular effect of each ESG dimension on the financial performance of SRI government bond portfolios, in line with previous studies on stock portfolios (Auer and Schuhmacher, 2016). Evaluating specific channels through which mutual funds could affect their holding bonds' social performance (Li et al. 2019). Different maturities of government bonds could be considered to form portfolios, as well as include more countries, especially developing countries, and evaluate a longer sample period. In addition, since previous research on SRI investment funds and SRI stock portfolios found that different market states (e.g., expansion and recession) affect the financial performance of SRI (Leite and Cortez, 2015; Gómez-Bezares et al. 2016), researchers could evaluate this concern about SRI government bond portfolios. Specific SRI issues assessed previously on firms could be analyzed from now on in this context.

References

Andrews, Donald WK; Monahan, J. Christopher. An Improved Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimator. *Econometrica: Journal of the Econometric Society*, 1992, 953-966.

Auer, Benjamin R. Do Socially Responsible Investment Policies Add Or Destroy European Stock Portfolio Value? *Journal of Business Ethics*, 2016, *135*, 381.

Auer, Benjamin R.; Schuhmacher, Frank. Do Socially (Ir) Responsible Investments Pay? New Evidence from International ESG Data. *The Quarterly Review of Economics and Finance*, 2016, *59*, 51-62.

Brammer, Stephen; Brooks, Chris; Pavelin, Stephen. Corporate Social Performance and Stock Returns: UK Evidence from Disaggregate Measures. *Financial Management*, 2006, *35*, 97-116.

Capelle-Blancard, G.; Crifo, P.; Diaye, M. A.;Scholtens, B.; Oueghlissi, R. Environmental, Social and Governance (ESG) Performance and Sovereign Bond Spreads: An Empirical Analysis of OECD Countries, 2016. Available online: https://ssrn.com/abstract=2874262.

Cortez, Maria Céu; Silva, Florinda; Areal, Nelson. Socially Responsible Investing in the Global Market: The Performance of US and European Funds. *International Journal of Finance & Economics*, 2012, *17*, 254-271.

Demiguel, Victor; Nogales, Francisco J. Portfolio Selection with Robust Estimation. *Operations Research*, 2009, *57*, 560-577.

Derwall, J; Guenster, N.; Bauer, R.; Koedijk, K. The Eco-Efficiency Premium Puzzle. *Financial Analysts Journal*, 2005, *61*, 51-63.

Derwall, Jeroen; Koedijk, Kees. Socially Responsible fixed-income Funds. *Journal of Business Finance & Accounting*, 2009, *36*, 210-229.

Directive 2014/95/EU of the European Parliament and of the Council of 22 October 2014 amending Directive 2013/34/EU as regards disclosure of non-financial and diversity information by certain large undertakings and groups

Drut, Bastien. Sovereign Bonds and Socially Responsible Investment. *Journal of Business Ethics*, 2010, *92*, 131-145.

Eccles, Robert G.; Ioannou, Ioannis; Serafeim, George. The Impact of Corporate Sustainability on Organizational Processes and Performance. *Management Science*, 2014, *60*, 2835-2857.

Eccles, Robert G.; Serafeim, George; Krzus, Michael P. Market Interest in Nonfinancial Information. *Journal of Applied Corporate Finance*, 2011, *23*, 113-127.

European Sustainable Investment Forum – EUROSIF (2016). European SRI Study 2016. Available on line: https://www.eurosif.org (accessed on 13 March 2018).

Ferruz, Luis; Sarto, J. Luis. An Analysis of Spanish Investment Fund Performance: Some Considerations Concerning Sharpe's Ratio. *Omega*, 2004, *32*, 273-284.

Galema, Rients; Plantinga, Auke; Scholtens, Bert. The Stocks at Stake: Return and Risk in Socially Responsible Investment. *Journal of Banking & Finance*, 2008, *32*, 2646-2654.

Gasbarro, Dominic; Wong, Wing-Keung; Kenton Zumwalt, J. Stochastic Dominance Analysis of iShares. *The European Journal of Finance*, 2007, *13*, 89-101.

Global Sustainable Investment Review (2016). Available online: https://www.gsi-alliance.org (accessed on 10 July 2018).

Goldreyer, Elizabeth F.; Diltz, J. David. The Performance of Socially Responsible Mutual Funds: Incorporating Sociopolitical Information in Portfolio Selection. *Managerial Finance*, 1999, *25*, 23-36.

Gómez-Bezares, Fernando; Przychodzen, Wojciech; Przychodzen, Justyna. Corporate sustainability and shareholder wealth—Evidence from British companies and lessons from the crisis. *Sustainability*, 2016, *8*, 276.

Gregory, Alan; Whittaker, Julie. Performance and Performance Persistence of 'ethical' unit Trusts in the UK. *Journal of Business Finance & Accounting*, 2007, *34*, 1327-1344.

Halbritter, Gerhard; Dorfleitner, Gregor. The Wages of Social responsibility—where are they? A Critical Review of ESG Investing. *Review of Financial Economics*, 2015, *26*, 25-35.

Hall, Peter. The Bootstrap and Edgeworth Expansion. 1992. Springer, New York, 1992.

Hoepner, Andreas GF; Neher, Agnes. Sovereign Debt and Sustainable Development Culture, 2013. Available online: https://ssrn.com/abstract=2295688

Hörisch, Jacob; Ortas, E.; Schaltegger, S; Álvarez, I. Environmental Effects of Sustainability Management Tools: An Empirical Analysis of Large Companies. *Ecological Economics*, 2015, *120*, 241-249.

Humphrey, Jacquelyn E.; Lee, Darren D.; Shen, Yaokan. Does it Cost to be Sustainable? *Journal of Corporate Finance*, 2012, *18*, 626-639.

Humphrey, Jacquelyn E.; Warren, Geoffrey J.; Boon, Junyan. What is Different about Socially Responsible Funds? A Holdings-Based Analysis. *Journal of Business Ethics*, 2016, *138*, 263-277.

Jobson, J. Dave; Korkie, Bob M. Performance Hypothesis Testing with the Sharpe and Treynor Measures. *The Journal of Finance*, 1981, *36*, 889-908.

Kempf, Alexander; Osthoff, Peer. The Effect of Socially Responsible Investing on Portfolio Performance. *European Financial Management*, 2007, *13*, 908-922.

Kreander, N.; Gray, R. H.; Power, D. M.; Sinclair, C. D. Evaluating the Performance of Ethical and non-ethical Funds: A Matched Pair Analysis. *Journal of Business Finance & Accounting*, 2005, *32*, 1465-1493.

Lahiri, SN. *Resampling Methods for Dependent Data.* Springer, New York., 2003.

Ledoit, Oliver; Wolf, Michael. Robust Performance Hypothesis Testing with the Sharpe Ratio. *Journal of Empirical Finance*, 2008, *15*, 850-859.

Leggio, Karyl B.; Lien, Donald. An Empirical Examination of the Effectiveness of Dollar-Cost Averaging using Downside Risk Performance Measures. *Journal of Economics and Finance*, 2003, *27*, 211.

Leite, Carlo, Cortez, Maria C., Silva, Florinda; Adcock, Christopher. The Performance of Socially Responsible Equity Mutual Funds: Evidence from Sweden. *Business Ethics: A European Review*, 2018, *27*, 108-126.

Leite, Paulo; Cortez, Maria Céu. Performance of European Socially Responsible Funds during Market Crises: Evidence from France. *International Review of Financial Analysis*, 2015, *40*, 132-141.

Leite, Paulo; Cortez, Maria Céu. The Performance of European Socially Responsible Fixed-Income Funds. *(January 8, 2016)*, 2016.

Li, Frank. Endogeneity in CEO power: A survey and experiment. *Investment Analysts Journal*, 2016, *45*, 149-162.

Li, Frank; Patel, Saurin; Ramani, Srikanth, The Role of Mutual Funds in Corporate Social Responsibility, **2019**. Available at SSRN: https://ssrn.com/abstract=3366100

Loh, Wei-Yin. Calibrating Confidence Coefficients. *Journal of the American Statistical Association*, 1987, *82*, 155-162.

Luo, Cuicui; Seco, Luis; Wu, Lin-Liang Bill. Portfolio Optimization in Hedge Funds by OGARCH and Markov Switching Model. *Omega*, 2015, *57*, 34-39.

Meligkotsidou, Loukia; Vrontos, Ioannis D.,Vrontos, Spyridon D. Quantile Regression Analysis of Hedge Fund Strategies. *Journal of Empirical Finance*, 2009, *16*, 264-279.

Memmel, Christoph. Performance Hypothesis Testing with the Sharpe Ratio. *Finance Letters, 2003, 1*.

Mollet, Janick Christian; Ziegler, Andreas. Socially Responsible Investing and Stock Performance: New Empirical Evidence for the US and European Stock Markets. *Review of Financial Economics*, 2014, *23*, 208-216.

Newey, Whitney K.; West, Kenneth D. A. Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation consistent Covariance Matrix, 1986.

Osthoff, P. The Routledge Handbook of Responsible Investment. Hebb, T., et al eds., 1st ed. Routledge, 2015. What Matters to SRI Investors? 705-724.

Politis, Dimitris N.; Romano, Joseph P. A Circular Block-Resampling Procedure for Stationary Data. *Exploring the Limits of Bootstrap*, 1992. pp. 263-270.

Revelli, Christophe; Viviani, Jean-Laurent. Financial Performance of Socially Responsible Investing (SRI): What have we Learned? A meta-analysis. *Business Ethics: A European Review*, 2015, *24*, 158-185.

Scholtens, Bert. A note on the interaction between corporate social responsibility and financial performance. *Ecological economics*, 2008, *68*, 46-55.

Scholz, Hendrik. Refinements to the Sharpe Ratio: Comparing Alternatives for Bear Markets. *Journal of Asset Management*, 2007, *7*, 347-357.

Sharpe, William F. Mutual Fund Performance. *The Journal of Business*, 1966, *39*, 119-138.

Sortino, Frank A.; Price, Lee N. Performance Measurement in a Downside Risk Framework. *The Journal of Investing*, 1994, *3*, 59-64.

Sortino, Frank A.; Van Der Meer, Robert. Downside Risk. *The Journal of Portfolio Management*, 1991, *17*, 27-31.

Statman, Meir; Glushkov, Denys. Classifying and Measuring the Performance of Socially Responsible Mutual Funds. *The Journal of Portfolio Management*, 2016, *42*, 140-151.

Statman, Meir; Glushkov, Denys. The Wages of Social Responsibility. *Financial Analysts Journal*, 2009, *65*, 33-46.

Ullmann, Arieh A. Data in Search of a Theory: A Critical Examination of the Relationships among Social Performance, Social Disclosure, and Economic Performance of US Firms. *Academy of Management Review*, 1985, *10*, 540-557.

Utz, Sabastian; Wimmer, Maximillian. Are they any Good at all? A Financial and Ethical Analysis of Socially Responsible Mutual Funds. *Journal of Asset Management*, 2014, *15*, 72-82.

Van De Velde, Eveline; Vermeir, Wimand Corten, Filip. Corporate Social Responsibility and Financial Performance. *Corporate Governance: The International Journal of Business in Society*, 2005, *5*, 129-138.

Appendix A. RobecoSAM information about ESG dimensions

Environmental dimension: Environmental challenges pose a potential risk for investors, as environmental externalities can result in significant economic losses, whereas repairing environmental damage, such as air and water pollution, can generate considerable fiscal costs. Adequate investments in preventing environmental problems limit such potential liabilities. Another important risk is related to the country's exposure to natural hazards, such as floods, hurricanes, or typhoons. In addition to evaluating a country's environmental vulnerabilities and policies, RobecoSAM examines its energy dependency and energy policies. Countries that rely heavily on fossil fuel imports are vulnerable to abrupt and/or sharp external price movements or supply shortages. In addition to assessing the risks themselves, RobecoSAM specifically looks for evidence that policies for mitigating such risks have been implemented.

Social dimension: A weak social climate dominated by labor unrest, extreme inequality, or other social tensions is another potential investment risk. A delicate social climate can easily result in violent turmoil, disrupting important economic activity, such as manufacturing or trade, and/or paralyze policymaking. Strong social cohesion, conversely, supports orderly conflict resolution and facilitates the implementation of necessary reforms, thus contributing to sustainable economic development.

Governance dimension: RobecoSAM examines a broad range of data that considers a country's institutional framework, regulatory quality, rule of law, government efficiency, central bank independence, and political stability, among other factors. Civil liberties, internal conflicts, and corruption also reflect a country's governance profile. The corruption level, for instance, shows the extent to which public power is exercised to protect the interests of a small group at the expense of the economy and society at large. A study by Robeco demonstrated the added value of considering political risk when taking investment decisions for government bonds, over a time period of 25 years.

Appendix B. Description of control variables

GDP growth rate: ΔGDP/GDP; annual percentages of constant-price GDP changes; source: International Monetary Fund (IMF).

Inflation: ΔP/P; annual percentages of average consumer price changes; source: IMF.

Fiscal Condition: Debt/GDP; all liabilities that require payment or payments of interest and/or principal by the debtor to the creditor at a date or dates in the future; source: IMF.

Primary Balance (PB): PB/GDP; primary net lending/borrowing plus net interest payable/paid; source: IMF.

Current Account (CA): CA/GDP; all transactions other than those in financial and capital items; source: IMF.

Liquidity ratio: Reserves/Imports; total reserves comprise holdings of monetary gold, special drawing rights, and holdings of foreign exchange under the control of monetary authorities; source: WB.

Country openness: (X + M)/GDP; the sum of exports and imports of goods and services measured as a share of gross domestic product; source WB.

Standard & Poor's (S&P) sovereign credit ratings: numerical variable assigning 1 to CCC, 2 to CCC+, and so on through 18 to AAA; source: Thomson Reuters.

Chapter 6: Are investments in material corporate social responsibly issues a key driver of financial performance?

Guillermo Badía[1], Fernando Gómez-Bezares[2], and Luis Ferruz[1]

[1]Faculty of Economics and Business, University of Zaragoza,
Zaragoza 50005, Spain.

[2]Deusto Business School, University of Deusto
48014 Bilbao, Spain.

Abstract

In this study, we address a lesser-studied aspect in corporate social responsibly (CSR): distinguishing between investments in material versus immaterial sustainability issues. This paper assesses the financial performance of stock portfolios formed according to material CSR issues and general CSR issues over the period 2007 to 2018. Our sample includes firms from Europe and the United States. Using several financial performance measures, we find that in Europe materiality is relevant to finding the best and worst firms both in terms of ESG and of financial performance whereas in the US market we do not find a significant positive effect of materiality on the firm's financial performance, although using the materiality the financial performance improves.

Keywords Socially responsible investing; Corporate social responsibility; Material sustainability issues; Portfolio performance evaluation

6.1 Introduction

As early as 1924, Sheldon (1924) introduced the concept of corporate social responsibility (CSR). Almost a hundred years later there is still an extensive ongoing debate about the relationship between CSR and corporate financial performance (CFP). Previous evidence so far is inconclusive with some studies finding a positive link between CSR and CFP, and others showing a negative or nonexistent relationship. Researchers have asserted that inconsistent findings in prior studies may be owing to neglecting factors such as making measurement errors, mis-specifying models, undersized and multi-industry samples, multi-dimensionality of CSR, or immaterial CSR investments (Wood and Jones, 1995; Khan et al., 2016; Man, 2017; Javed et al., 2016). Despite conflicting results, institutional and private investors have gradually incorporated socially responsible firms into their investment portfolios (Ferruz et al., 2012). They evaluate not only financial criteria (returns and risk) in their investment decisions, but rather the non-financial attributes of socially responsible investing (SRI) (Galema et al., 2008). Environmental, social and governance (ESG) issues are becoming more important in investors' decision-making in the efforts to help identify the long-term opportunities and risks for firms. Rising individual awareness of environmental, social and ethical issues is strongly influencing purchase decisions of investors (Mollet and Ziegler, 2014). Asset managers considered ESG criteria across $11.6 trillion in assets, having increased by 44% from $8.1 trillion in 2016 (USSIF, 2018). The EUROSIF (2018) report discloses sustained growth for the most sustainable and responsible investment strategies. The past two years (2016-2018) have shown signs of SRI becoming integral to European fund management.

Many firms are also paying closer attention to CSR policies, strategies and demands. The positive relationship between CSR and CFP may be attributable to the fact that only successful firms have their sources to engage in CSR-related activities (Ullmann, 1985). However, CSR may positively influence CFP in several ways. In fact, several literature review studies document a positive effect of CSR on CFP and valuation (e.g., Orlitzky et al., 2003; Margolis et al., 2009; Lu et al., 2014; Javed et al., 2016; and Man, 2017). Firms enhancing, for instance, employee incentives may achieve that higher levels of motivation among employees do better, increasing their productivity and resulting in more efficient manufacturing processes. Therefore, improving incentives will have positive effects on corporate products, brand, reputation, and consequently sales will increase. Moreover, firms improving, for instance, manufacturing processes using

technology may reduce emissions, time-processes, work accidents, and enhance employees' safety, resulting in lower costs resulting from safety issues, for example, infractions or penalties. This could result in better community relations and could help firms to obtain alternative and additional financial resources from socially conscientious intermediaries.

In this study, we address a lesser-studied aspect in CSR: distinguishing between investments in material versus immaterial sustainability issues. Khan et al. (2016) highlight that one potential reason for the inconclusive results on the link between CSR and CFP is that previous studies do not distinguish between sustainability issues that are material for a firm versus immaterial sustainability issues. Material aspects refer to issues that, when managed effectively, represent a significant contribution to the firm's value, but if not, may lead to a significant loss of value and opportunities to create or preserve future value (Eccles and Youmans, 2016). Firms paying attention to sustainability issues that are associated with their main operations and focused on enhancing in the material areas where their performance is lower, will gain a competitive advantage over their competitors and achieve a higher corporate social and financial performance (Khan et al., 2016).

Material issues are those which have a significant financial impact for firms in a particular industry. Herz and Rogers (2016) note the importance of using a targeted approach to rating ESG practices of firms based on material items for each industry. For instance, for automobile firms, investors want to evaluate progress on developing alternative-fuel vehicles to curb use-phase emissions and capitalize on changing consumer preferences. In the case of commercial banks, investors want to know about financed emissions –loans to oil and gas companies, and to industrial and utility companies. For software and IT companies, investors want to know the energy intensity of data centers, which carries regulatory and reputational risks along with innovation opportunities.

According to the *2019 ESG Trends to Watch* report (MSCI, 2019), investors now have to turn their attention from data proliferation to relevant signals. Although there is a lot of ESG and sustainability information disclosed publicly, often it is difficult to identify and assess which information is most useful for making financial decisions. Investors currently become familiar with the ESG data framework, but the most successful will be those who recognize that they have an advantage only if they have a clear view of a material signal. From now on, having more data will be the easy part; the hard, and

material, part will be identifying the most relevant signal to achieve an improved financial performance. The Sustainability Accounting Standards Board (SASB) developed the Materiality Map to help investors with this issue. It allows both investors and firms to identify financially material issues across different industries, which are reasonably likely to impact the financial condition or operating performance of a firm, and therefore are most important. Materiality is relevant to firms so they can focus their sustainability strategies on the most important issues; for investors, materiality is important in evaluating portfolio exposure to specific material and immaterial sustainability risks and opportunities.

This paper assesses the financial performance of stock portfolios formed according to material CSR issues and general CSR issues over the period 2007 to 2018. Our main research question is: Are investments in material CSR issues a key driver of financial performance? Khan et al. (2016) showed that US firms with strong performance on material aspects outperformed firms with poor performance on material topics. Our dataset includes companies from United States and Europe. Evaluating firms from United States and Europe is particularly interesting given the heterogeneity in the patterns of development of SRI across countries (Neher and Hebb, 2015). The extension of SRI research to other geographical areas is further motivated by Hörisch et al. (2015), who indicated that country-specific factors tend to affect the relationship between corporate social and financial performance. Investors' ESG concerns can differ from region to region. For instance, Eccles et al. (2011) found that European investors are more concerned with environmental information, while US investors are more interested in governance issues. In turn, Cortez et al. (2012) identified geographical differences in the investment style of socially responsible funds. We use firm's scores from an original dataset, Truvalue Labs, which, as far as we are aware, has not been used before in this regard. Truvalue Labs has integrated the SASB Materiality Map standards into Truvalue Labs dataset. Truvalue Labs collects and analyzes information related to the leading industry standards set by SASB for ESG factors that are material for financial performance in each industry. Data is collected from more than 100,000 sources to provide insights and analysis for positive and negative portfolio filtering and company monitoring.

The structure of the paper is as follows: Section 2 presents an overview on the effects of CSR on financial performance and provides the hypotheses development. Section 3

describes the data and Section 4 contains the empirical analysis used. Section 5 summarizes and discusses our main findings and offers some concluding remarks.

6.2 CSR insights and hypothesis

Many studies have hypothesized on the effects of CSR strategies on firm stakeholders. Man (2017) highlighted that CSR affects all aspects of firms, both internal corporate operations and behavior of external stakeholders. According to Armstrong and Green (2013) stakeholders are creditors, costumers, distributors, employees, local communities, suppliers, owners; i.e., 'any group of individuals who can affect or is affected by the achievement of a firm's objectives' (Freeman, 1984). Obviously, employee motivations and rights are a key aspect of the economic health of firms. Companies enhancing employee relations can, for example, encourage employees to be more efficient, take greater care in their relationships with clients and suppliers, accept voluntary work, and, in the end, align employee motivations with corporate goals, therefore increasing employees productivity and the firm's economic performance. Sharing goals with suppliers and knowing their needs also seems relevant to establishing efficient supply chains and therefore avoiding problems with provisions or manufacturing processes. It allows firms to offer corporate products to clients on time. On the other hand, CSR strategies may provide firms with a reduction of corporate risks related to environmental concerns. Firms concerned with environmental aspects are better equipped to deal with environmental requirements and to innovate on cleaner manufacturing processes. Innovation on social and environmental aspects allows firms to search and achieve joint solutions to problems linked to stakeholders. Open innovation creates new solutions calling for significant stakeholder interaction to achieve them (Chesbrough 2003). Solving problems in society demands a constant collaboration among all actors, and social innovation is crucial to transform an idea into a solution that creates value for stakeholders (Osburg, 2013). By improving CSR strategies and showing them to stakeholders, firms can enhance the reputation associated with the brand and increase their financial performance. Reputation is a bottom determinant in the relationship between CSR and financial performance of firms. Reputation benefits consumer perceptions and tends to decrease consumers' price sensitivity and increase their brand loyalty.

Previous empirical studies evaluate these arguments. For example, Filbeck and Preece (2003), Fulmer et al. (2003), and Edmans (2011) assessed the relationship between

employee satisfaction and financial performance of firms using stocks listed in the '100 Best Companies to Work For in America'. They find that companies with stronger employee satisfaction outperform conventional firms. Derwall et al. (2005) compared the financial performance of two stock portfolios that differ in eco-efficiency characteristics and found that a portfolio of high-ranked eco-efficiency firms outperforms a portfolio of low-ranked ones. Filbeck et al. (2013) assessed whether the fact of being listed on public surveys of exceptional companies (for instance *Business Ethics* 'Best Corporate Citizens') adds value to a portfolio. They found that firms on the 'Most Admired Companies' and the 'Best Corporate Citizens' rankings are the most influential ones. Despite these results, empirical evidence on benefits of CSR activities on the financial performance of firms is still far off consensus (Badía et al., 2018). One potential reason for the inconsistent results is the fact that the research does not distinguish between material and immaterial sustainability issues (Khan et al., 2016). However, only firms focused on material sustainability issues associated with their main operations will gain a competitive advantage and achieve a higher social and financial performance. CSR activities and innovations should be performed on material aspects, because otherwise a positive effect on financial performance is not expected. Indeed, investments on immaterial issues may involve additional corporate costs without an associated social and financial performance return. Focusing on material issues is important for firms since they invest in social aspects that profoundly affect their operations. Despite the fact that issues such as product safety, climate change, and resource use intensity have an impact across several industries, as Herz and Rogers (2016) noted, those effects often vary to a great extent from one industry to the next. Risks may be everywhere, although they are indeed also specific. Consequently, firms in specific industries have their own particular sustainability profiles. Therefore, a firm investing and reporting on material sustainability issues is likely to achieve positive financial performance. Meanwhile, a firm investing in material but also in immaterial sustainability issues is not likely to achieve superior financial performance. These arguments lead to our two first hypotheses:

Hypothesis 1: Firms scoring high on material issues outperform firms scoring low on material issues.

Hypothesis 2: Firms scoring high on material and immaterial issues perform similar to firms scoring low on material and immaterial issues.

In this paper, we evaluate firms from North America and Europe. This assessment is of particular interest given the heterogeneity in the patterns of development of CSR and the different SRI strategies implemented on these regions (Neher and Hebb, 2015). Louche and Lydenberg (2006) explored the development and main practices of CSR in the US and EU markets, and showed that there are differences in terms of definitions of CSR, SRI screening strategies, involved actors leading CSR role, and approaches to engagement with firms by the CSR community. For example, they noted that in Europe, environmental issues hold greater emphasis than in the US. Eccles et al. (2011) also identified the fact that European investors are more interested in environmental aspects, while US investors are more concerned with governance issues. In fact, they noted that in the US market there is considerable skepticism about the potential effects of climate change. The findings of Badía et al. (2018) confirmed the regional and cultural idiosyncrasies in SRI. They showed that North American firms are more sensitive to governance issues, while European firms paid closer attention to social and environmental demands. Sandberg et al. (2009) suggested that there are at least three cultural and ideological differences between different regions: differences in values and norms, ideology among different SRI stakeholders, and the market setting of SRI. Louche and Lydenberg (2006) argued that, indeed, differences in definitions reflect cultural differences between the US and EU markets. Despite the differences between the two regions, they both place emphasis on using the investment process as a means to change and improve the behavior of firms on social and environmental issues. In fact, Ullmann (1985) noted that investors have the power to influence management's CSR activities and SRI demands have led firms to be more concerned with their CSR strategies.

These cultural and ideological differences between regions likely affect the financial performance of the firms. In fact, empirical studies evaluating the financial performance of firms from the US and EU found distinct outcomes. Auer and Schuhmacher (2016), for instance, found that high-ranked US firms according to ESG dimensions perform similarly to low-ranked ones, while high-ranked European firms, depending on the industry and the ESG screening used, perform worse than low-ranked firms. Badía et al. (2017), evaluating the relative financial performance between SRI North American firms and SRI European firms, found that North American firms outperform their European counterparts. Nevertheless, studies on the specific markets show mixed results. For example, in the US market, Borgers et al. (2013), and Eccles et al. (2014)

found that high-ranked firms according to ESG dimensions outperform low-ranked ones, but Lee et al. (2013), and Halbritter and Dorfleitner (2015) did not find significant financial differences between high- and low-ranked sustainable firms. In Europe, Mollet et al. (2013) and Auer (2016) found that high-ranked firms outperform low-ranked ones, but Van de Velde et al. (2005), and Humphrey et al. (2012) did not find any significant differences between high- and low-ranked firms.

Both social and financial performance differences identified in previous research lead us to presume that firms from North America and Europe may perform differently. Thus, our third hypothesis is established as follows:

Hypothesis 3: Firms from North America perform differently than firms from Europe in terms of financial performance associated to CSR aspects.

6.3 Data

Our sample includes firms from Europe and the United States over the period of 2007 to 2018. We evaluate firms from the main stock exchanges of both regions with social responsibility scores[25]. Monthly discrete returns of all stocks are computed based on the total return series (in US dollars) collected from the Thomson Reuters database. We use the social responsibility ratings of companies provided by the Truvalue Labs database to form portfolios. Truvalue Labs provides an overall score and sustainability trend, as well as specific performance for individual categories. Truvalue Labs delivers timely material ESG Insights using the SASB standards, which are widely considered the industry standard for identifying material sustainability issues by industry. Truvalue Labs has integrated the SASB Materiality Map standards into Truvalue Labs. The SASB's Materiality Map standards includes 5 sustainability dimensions and 30 general issue categories, which make up SASB's universe of ESG issues. Underlying each general issue category is a set of industry-specific sustainability disclosure topics in the SASB standards. These categories are defined by the SASB, and the categories considered material to financial performance vary by industry. Truvalue Labs is independent but it is partnered with the SASB to develop Truvalue Labs SASB Edition. The SASB Edition uses standards from the industry-leading Sustainability Accounting Standards Board. Table 1 displays the evolution of the number of stocks included in the sample. As expected, the number of stocks increases throughout the sample period. Also, firms from the US market are more evaluated than firms from the EU market,

[25] Appendix A displays stock exchanges where firms are traded.

likely due to the fact that the SASB was founded in the US seeking to create industry sustainability standards for the disclosure and recognition of financially material environmental, social, and governance impacts of publicly traded US firms.

Table 6-1. Stocks over the sample period (2007-2018)

This table presents the region stocks allocation each year from US and EU markets. The full sample period is from January 2007 to December 2018.

Period	US	EU
2007	118	27
2008	735	199
2009	1058	257
2010	1291	302
2011	1569	351
2012	1796	368
2013	2008	399
2014	2223	440
2015	2558	474
2016	2881	532
2017	2842	560
2018	2853	573

6.4 Empirical analysis

6.4.1 Portfolio formation

We form portfolios using the materiality score which aggregates only the material categories, as defined by the SASB, and also using the all-category score which aggregates all 30 categories as defined by the SASB. The Materiality portfolios are formed each year as equally-weighted portfolios of firms' stocks based on their materiality scores in the previous year[26]. The high-rated portfolio comprises stocks with the best materiality rated firms and the low-rated portfolio includes those with the worst materiality rated firms. The All-category portfolios are constructed in the same way. As in prior studies (e.g., Van de Velde et al., 2005; Kempf and Osthoff, 2007; Derwall et al., 2011; Halbritter and Dorfleitner, 2015; and Auer, 2016), we use different cut-offs to form the portfolios (10%, 20%, and 30%), thus allowing us to evaluate portfolios that are more restricted or broader with respect to the social criteria used. Then, we form the difference portfolio, which is obtained by subtracting the low-ranked portfolio returns from the returns on the high-ranked stock portfolio, thus representing a strategy of going long in the high-rated stocks and short in the low-rated stocks. The analysis of the performance of the long-short portfolios enables us to conclude whether there are

[26]Truevalue Labs performs a daily update of the data, ie, the ESG scores of firms at day *t-1* are updated at day *t* around 5am.

statistically significant differences between the performance of high- and low-rated portfolios.

Table 2 presents descriptive statistics of the Materiality and All-category portfolios under different cut-offs[27].Comparing for the different cut-offs, the high-rated portfolios show higher average returns than the low-rated ones in both markets using materiality and all-category information. We observe that the biggest differences appear using the materiality scores in both markets under the most demanding level, the 10% cut-off. This can be considered, initially, as evidence that materiality could be relevant to the financial performance of portfolios. As for standard deviation, in the EU market, the high-ranked portfolios show, in general, a similar variability of returns than low-ranked ones, whereas in the US market, the high-ranked portfolios show a higher variability than low-ranked ones. These findings encourage the use of risk-adjusted measures to evaluate the financial performance of portfolios.

Table 6-2. Descriptive statistics

This table displays a summary statistic of the high- and low-rated portfolios at the 10% [10], 20% [20], and 30% [30] cut-offs. Mean (SD) is the average return (standard deviation) of portfolios. Diff is the mean (SD) difference between high- and low-rated portfolios. The full sample period is from January 2007 to December 2018.

Panel A. EU market						
Materiality	High [10]	Low [10]	High [20]	Low [20]	High [30]	Low [30]
Mean	0.0066	0.0023	0.0066	0.0039	0.0059	0.0042
Diff	0.0043		0.0027		0.0017	
SD	0.0596	0.0600	0.0623	0.0604	0.0624	0.0621
Diff	-0.0004		0.0019		0.0002	
All Categories	High [10]	Low [10]	High [20]	Low [20]	High [30]	Low [30]
Mean	0.0075	0.0045	0.0068	0.0038	0.0064	0.0042
Diff	0.0031		0.0029		0.0022	
SD	0.0616	0.0605	0.0619	0.0611	0.0622	0.0622
Diff	0.0012		0.0008		0.0000	
Panel B. US market						
Materiality	High [10]	Low [10]	High [20]	Low [20]	High [30]	Low [30]
Mean	0.0059	0.0040	0.0052	0.0041	0.0048	0.0046
Diff	0.0019		0.0011		0.0002	
SD	0.0594	0.0569	0.0584	0.0541	0.0580	0.0543
Diff	0.0025		0.0042		0.0037	
All Categories	High [10]	Low [10]	High [20]	Low [20]	High [30]	Low [30]
Mean	0.0058	0.0057	0.0055	0.0049	0.0056	0.0051
Diff	0.0001		0.0006		0.0005	
SD	0.0577	0.0551	0.0573	0.0552	0.0562	0.0546
Diff	0.0026		0.0021		0.0015	

6.4.2 Financial performance

Portfolio financial performance is evaluated using several approaches, in line with Scholtens (2008), and Carvalho and Areal (2016). The magnitude and sometimes even

[27]Appendix B displays descriptive statistics in-depth.

the sign of the long-run abnormal returns are sensitive to alternative measurement methodologies (Fama, 1998; Loughran and Ritter, 2000).

6.4.2.1 Penalized Internal Rate of Return (PIRR)

To start, we use the PIRR ratio given by the following equation (Gómez-Bezares and Gómez-Bezares, 2012):

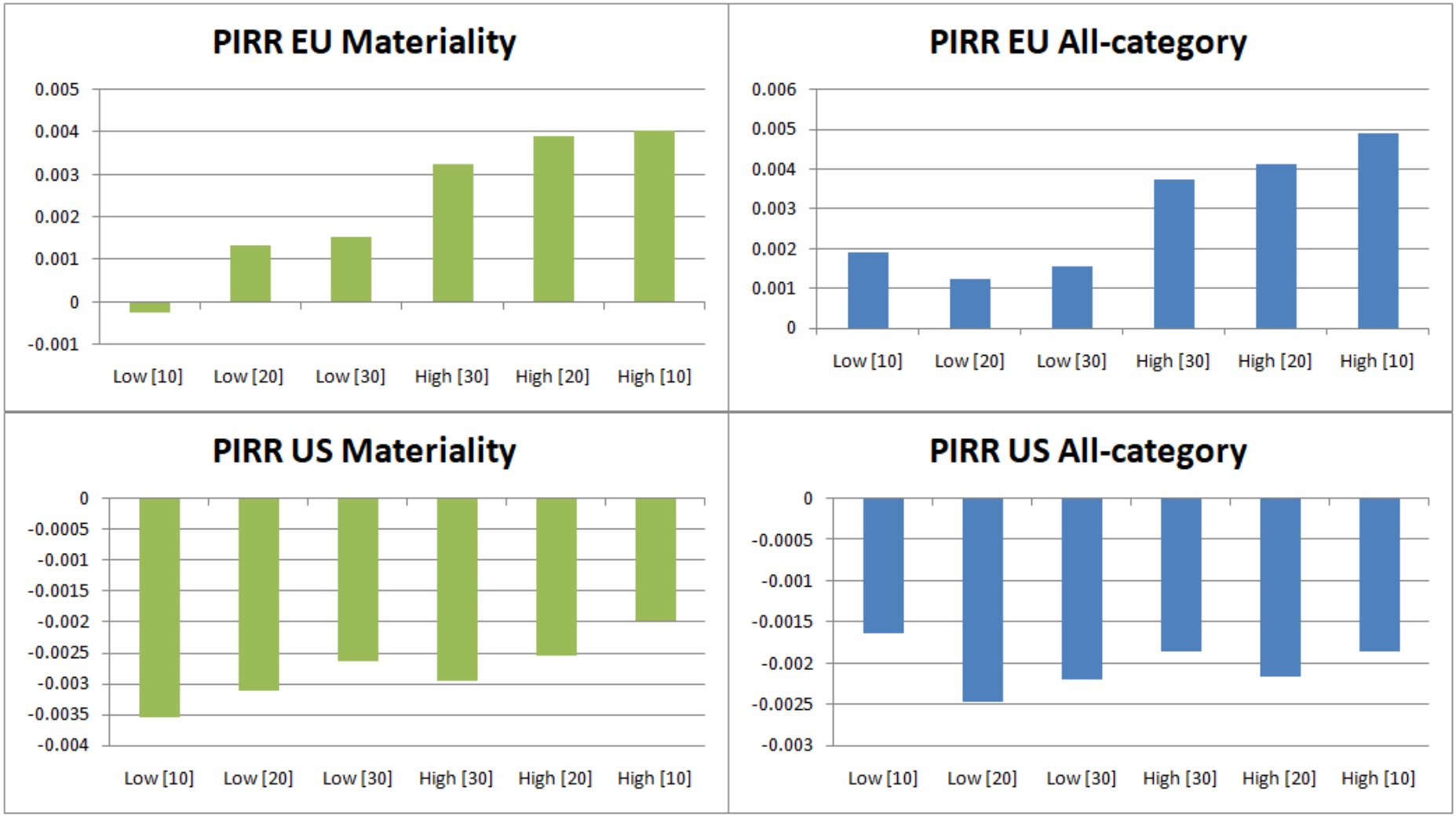
$$PIRR_T = \mu_T - \left[\frac{(\mu_{mt} - r_{ft})}{\sigma_{mt}}\right] \times \sigma_t$$

(Eq. 6-1)

where μ_T is the average monthly return on a given portfolio at year *t*, μ_{mt} is the average monthly return on the market portfolio at year *t*, r_{ft} is the monthly return on a risk-free asset for year *t*, σ_{mt} is the standard deviation of the rate of return on the market portfolio for year *t*, and σ_t is the standard deviation of the rate of return on the given portfolio for year *t*. The market portfolio is specific for each region and they are obtained from Professor Kenneth French's website. In line with Gómez-Bezares et al. (2016), we interpret $PIRR_T$ as the reward-to-variability performance measure for total risk.

The results of applying the PIRR ratio are presented graphically. Figure 1 shows the PIRR values associated with the Materiality and All-category portfolios at the different cut-offs in the EU and US markets. In the EU market, using materiality scores we find that the more restricted the cut-off used is–from 30% to 10%– the better the high-rated portfolios perform, whereas the low-rated portfolios perform worse. These results suggest that investors who are more socially demanding, since the firms included in their portfolios are the best in material issues, achieve better financial performance. This is in line with the claim of Khan et al. (2016), who noted that firms focused on material sustainability issues will gain a competitive advantage over their competitors and achieve a higher corporate social and financial performance. Using all-category scores, we find a similar behaviour but with some differences. The low-rated portfolio at the 10% cut-off achieves better financial performance than those of the low-rated portfolios at the 20% and 30% cut-offs. This suggests that all-category scores are not as discriminatory as materiality scores to selecting firms that are, not only the best or worst in socially responsible aspects, but also in financial performance.

Figure 6-1. PIRR values associated with the Materiality and All-category portfolios at the different cut-offs in the EU and US markets



In the US market, using materiality scores we find a view similar to that in the EU market. The more restricted the cut-off used is, the better the high-rated portfolios perform, whereas the low-rated portfolios perform worse, except in the less socially responsible demanding level: the 30% cut-off. Using all-category scores, we find quite a different result. The low-rated portfolio performs better than the high-rated portfolio at the 10% cut-off. This evidence suggests that all-category scores do not allow identifying those firms performing high in socially responsible issues and also in financial performance. In sum, these findings suggest that the materiality scores are more useful than the all-category scores in selecting the best firms in terms of social aspects and of financial performance.

6.4.2.2 Multi-factor model

To evaluate portfolio performance, we also compute alphas from a multi-factor model, as for example in Van de Velde et al. (2005), Edmans (2011), Humphrey et al. (2012), and Badía et al. (2017). We examine performance using the four-factor Carhart (1997) model that allow capturing the risk premiums associated with size and value/growth (as in Fama and French, 1993) as well as momentum, motivated by Jegadeesh and Titman (1993). The Carhart (1997) four-factor model is expressed by:

$$R_{p,t} - R_{f,t} = \alpha_p + \beta_{RMRF} RMRF_t + \beta_{SMB} SMB_t + \beta_{HML} HML_t + \beta_{MOM} MOM_t + \varepsilon_{p,t} \quad \text{(Eq. 6-2)}$$

where $R_{p,t}$ is the return of portfolio p at time t, $R_{f,t}$ is the risk-free rate and α_p is the estimated performance measure of the portfolio. In relation to the risk factors, $RMRF_t$ represents market excess returns (relative to the risk-free rate) at time t; SMB_t is the difference between the returns on diversified portfolios of small stocks and large stocks; HML_t is the difference between the returns on diversified portfolios of high book-to-market (value) stocks and low book-to-market (growth) stocks; and MOM_t is the difference between the returns on diversified portfolios of winning and losing stocks in the past year. The betas in the model represent the estimated risk measures associated to the risk factors: market, size, value/growth and momentum. Finally, $\varepsilon_{p,t}$ represents the residuals. The specific independent variables for each region are obtained from Professor Kenneth French's website.

Table 3 displays the financial performance of the high- and low-rated portfolios as well as the long-short (L-S) portfolios[28]. Statistic alphas of the high- and low-rated portfolios are associated with outperforming the market benchmarks, i.e. conventional investments, as for instance in Carvalho and Areal (2016). In the EU market, both the Materiality and the All-category high-rated portfolios obtain positive and significant extra financial performance -at the 1% significance level- using any cut-off. Furthermore, some low-rated portfolios obtain positive and significant alphas although to lower levels. These results suggest that high-rated portfolios outperform conventional investments and that some low-rated ones do it as well. However, given the only positive and statistically significant alpha of the Materiality long-short portfolio at the 10% cut-off, only materiality scores allow us to identify the best and worst firms both in terms of ESG and of financial performance. This evidence suggests that, in Europe, firms investing well in material issues related to their industries, and thereby obtaining a high materiality score, are capable of outperforming firms investing poorly in material issues. In the US market, most high- and low-rated portfolios obtain negative financial performance, although not significant. Material issues, in this case, do not allow us to identify the best firms in terms of ESG and financial performance, but they do allow us to identify substantial negative financial performance associated to some low-ranked firms. For example, the low-rated portfolio at the 20% cut-off obtains a negative and statistically significant alpha, and the low-rated portfolio at the 10% cut-off obtains a negative alpha of 0.19%. Consequently, using the materiality scores, the long-short portfolios at the 10% and 20% cut-off achieve positive financial performance. These results are in line with our earlier results using the PIRR ratio. Despite the fact that in many cases no significant alphas associated to the long-short portfolio are found, results by PIRR show an clear charted pattern in portfolio performance behaviour. In general, using materiality scores, the more restricted the cut-off used, the better high-rated portfolios perform whereas the low-rated portfolios perform worse. However, using all-category scores, we cannot identify a tendency as evident as this. In sum, these findings confirm our first hypothesis graphically on US firms and graphically and statistically on EU firms, our second hypothesis for EU and US firms, and also our third hypothesis in terms of financial performance linked to material and immaterial issues.

[28]As we are focusing on the performance of SRI portfolios, only the alphas of the portfolios are reported. Nonetheless, coefficients related to specific risk-factors are available upon request.

Table 6-3. Portfolio financial performance

This table shows estimates of alpha (abnormal returns) of the high- and low-rated portfolios as well as the long-short portfolios at the 10% [10], 20% [20], and 30% [30] cut-offs using materiality and all-category scores. The long-short portfolio is formed by subtracting the returns of the low-ranked portfolio from the returns of the high-ranked portfolio (L-S). Panel A displays results for the EU market and Panel B shows results for the US market. Portfolio performance is evaluated by means of the alpha from the four-factor Carhart (1997) model. The model is estimated by OLS based on the heteroskedasticity and autocorrelation adjusted errors of Newey and West (1987). The independent variables are obtained from Professor Kenneth French's website. The asterisks are used to represent the statistically significant coefficients at the 1% (***), 5% (**) and 10% (*) significance levels. The full sample period is from January 2007 to December 2018.

Panel A. EU market									
Materiality	High [10]	Low [10]	L-S [10]	High [20]	Low [20]	L-S [20]	High [30]	Low [30]	L-S [30]
Alpha	0.0046***	0.0012	0.0034**	0.0044***	0.0027*	0.0016	0.0039***	0.0030**	0.0009
t-stat	2.9383	0.6843	2.3067	3.4835	1.7140	1.3631	3.7219	2.0731	0.7536
All Categories	High [10]	Low [10]	L-S [10]	High [20]	Low [20]	L-S [20]	High [30]	Low [30]	L-S [30]
Alpha	0.0061***	0.0032	0.0029	0.0052***	0.0027*	0.0024*	0.0046***	0.0029*	0.0017
t-stat	5.2668	1.5141	1.4215	5.2419	1.7545	1.9246	4.5184	1.7366	1.3463
Panel B. US market									
Materiality	High [10]	Low [10]	L-S [10]	High [20]	Low [20]	L-S [20]	High [30]	Low [30]	L-S [30]
Alpha	-0.0001	-0.0019	0.0017	-0.0008	-0.0015*	0.0007	-0.0013	-0.0010	-0.0003
t-stat	-0.0823	-1.5914	1.1162	-0.6664	-1.6777	0.6401	-1.3786	-1.1476	-0.4291
All Categories	High [10]	Low [10]	L-S [10]	High [20]	Low [20]	L-S [20]	High [30]	Low [30]	L-S [30]
Alpha	0.0000	0.0005	-0.0005	-0.0004	-0.0005	0.0000	-0.0002	-0.0003	0.0001
t-stat	-0.0387	0.4454	-0.4233	-0.4528	-0.5276	0.0484	-0.2401	-0.3719	0.1475

6.5 Discussion and conclusions

Our results for the EU market are new. We show that materiality is relevant to finding the best and worst firms both in terms of ESG and of financial performance. Furthermore, we find this effect under the most SRI demanding level, the 10% cut-off. As for the findings about the US market, our results contrast to Khan et al. (2016) since they find that US firms with strong performance on material aspects outperformed firms with poor performance on material topics, and we do not find a significant positive effect of materiality on the firm's financial performance. Nevertheless, it is relevant that we find that material aspects allow for the identification of the worst firms in terms of ESG and financial performance. The different results from Khan et al. (2016) could be due to several reasons. Firstly, in order to evaluate the implications of sustainability investments on the financial performance, they orthogonalise a firm's change in materiality score with respect to changes in firm size, market-to-book ratio, leverage, profitability, and sector membership. This process attempts to isolate unexplained changes in the score of firms. We implemented this process -orthogonalisation- on the materiality and all-category scores of each firm and our results are unchanged.

Secondly, the sample periods are different. Khan et al. (2016) assess the period from 1993 to 2013, whereas we evaluate the period from 2007 to 2018. Both periods share bull and bear times but, since their period is longer than ours, turbulent market states,

such as the international financial crisis from 2008, could be offset by up periods, affecting the results on the financial performance of portfolios over the full sample period. Our data on materiality scores of firms start in 2007 so we are limited to investigate this period. It would be of interest to extend the analysis to a longer period in order to assess whether it affects the results. In this regard, since previous studies identify that different market states (e.g., bull and bear periods) affect the financial performance of SRI portfolios (e.g., Becchetti et al., 2015; Leite and Cortez, 2015; Carvalho and Areal, 2016), we investigated whether market phases are affecting the financial performance of our portfolios based on material scores. We follow Nofsinger and Varma (2014) and Badía et al. (2018) and include two dummy variables in the previously used four-factor model. The model allows risk and performance to vary across different market states. To identify the different market states across our sample period we use the Pagan and Sossounov (2003) approach, in line with Lee et al. (2013) and Badía et al. (2017). The results show that US high- and low-ranked firms according to materiality scores perform similarly in bull and bear periods. These findings suggest that bull and bear phases do not affect the financial performance of US ranked firms according their materiality scores throughout our sample period, thereby ruling out a market state effect.

Thirdly, materiality scores are constructed differently. Khan et al. (2016) based their scores on the materiality guidance from the SASB using MSCI KLD as a source of sustainability data. A hand process is done by them to classify each KLD item as material, thereby may run into potential selection bias trouble. Specifically, they download each industry standard which identifies material sustainability issues for firms within an industry. To classify topics, one researcher took the lead in one sector and all the industries included in that sector. Each topic identified by the SASB as material was mapped to a KLD item, when one is available. After having completed the map, another researcher followed the same process. The two maps were then compared by a third researcher, who assessed any differences. As we detail in Data section, the process of identifying material issues by TrueValue Labs is substantially different, and as a result, our results could also be substantially different. The use of different ESG information sources has been one of the most highlighted factors causing different results on the financial performance of SRI (Javed et al., 2016; Henriksson et al. 2018). For instance, Halbritter and Dorfleitner (2015) find that the overall ESG scores of ASSET4 and Bloomberg both have a significant influence on the returns. However, the overall KLD

scores do not provide evidence for a link between the ESG level and the financial performance. Mixed results are, for example, Derwall et al. (2005) who found that US high-ranked firms according to Innovest ESG scores outperform low-ranked firms whereas neither Galema et al. (2008) using KLD scores, nor Lee et al. (2013) using SAM scores, found significant differences. As for materiality, Eccles and Youmans (2016) noted that materiality, in its essence, is entity-specific. Material interests and issues of the stakeholders change from firm to firm, depending on sector, strategy, business model, and the time frame under consideration. We consider that materiality is relevant to both firms and investors since it allows firms to focus their sustainability strategies on the most important issues and it allows investors to evaluate portfolio exposure to specific material and immaterial sustainability risks and opportunities. However, the material entity-specific information still seems to be inaccurate because, in the US market, materiality scores do not enable us to differentiate between the best firms and the worst, both in terms of ESG and of financial performance. Hence, we consider that material issues should still be defined in a more appropriate manner so that investors and firms can use them as an appropriate tool.

References

Armstrong, J. S., & Green, K. C. (2013). Effects of corporate social responsibility and irresponsibility policies. *Journal of Business Research*, *66*(10), 1922-1927.

Auer, B. R. (2016). Do socially responsible investment policies add or destroy European stock portfolio value? Journal of Business Ethics, 135(2), 381-397.

Auer, B. R., & Schuhmacher, F. (2016). Do socially (ir) responsible investments pay? new evidence from international ESG data. The Quarterly Review of Economics and Finance, 59, 51-62.

Badía, G., Cortez, M., & Ferruz, L. (2018). *The performance of socially responsible stock portfolios: International evidence.* Working paper.

Badía, G., Ferruz, L., and Cortez, M. (2017, September). SRI to retail investors in international stock markets and on different market states. XIX AECA International Conference, Santiago de Compostela, Spain.

Becchetti, L., Ciciretti, R., Dalò, A., & Herzel, S. (2015). Socially responsible and conventional investment funds: Performance comparison and the global financial crisis. Applied Economics, 47(25), 2541-2562.

Borgers, A., Derwall, J., Koedijk, K., & Ter Horst, J. (2013). Stakeholder relations and stock returns: On errors in investors' expectations and learning. Journal of Empirical Finance, 22, 159-175.

Carhart, M. (1997). On persistence in mutual fund performance. The Journal of Finance, 52(1), 57-82.

Carvalho, A., & Areal, N. (2016). Great places to work®: Resilience in times of crisis. Human Resource Management, 55(3), 479-498.

Chesbrough H (2003) Open innovation: the new imperative for creating and profiting from technology. Harvard Business School Publishing, Boston

Cortez, M. C., Silva, F., & Areal, N. (2012). Socially responsible investing in the global market: The performance of US and European funds. International Journal of Finance & Economics, 17(3), 254-271.

Derwall, J., Guenster, N., Bauer, R., & Koedijk, K. (2005). The Eco-Efficiency Premium Puzzle. Financial Analysts Journal, 61(2), 51-63.

Derwall, J., Koedijk, K., & Ter Horst, J. (2011). A tale of values-driven and profit-seeking social investors. Journal of Banking & Finance, 35(8), 2137-2147.

Eccles, R. G., & Youmans, T. (2016). Materiality in corporate governance: the statement of significant audiences and materiality. *Journal of Applied Corporate Finance*, *28*(2), 39-46.

Eccles, R. G., Ioannou, I., & Serafeim, G. (2014). The impact of corporate sustainability on organizational processes and performance. Management Science, 60(11), 2835-2857.

Eccles, R. G., Serafeim, G., & Krzus, M. P. (2011). Market interest in nonfinancial information. Journal of Applied Corporate Finance, 23(4), 113-127.

Edmans, A. (2011). Does the stock market fully value intangibles? employee satisfaction and equity prices. *Journal of Financial Economics*, *101*(3), 621-640.

European Sustainable Investment Forum – EUROSIF (2018). European SRI Study 2018. Available online: https://www.eurosif.org (accessed on 13 February 2019).

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. Journal of Financial Economics, 33(1), 3-56.

Fama, E. F. (1998). Market efficiency, long-term returns, and behavioral finance. *Journal of financial economics*, *49*(3), 283-306.

Ferruz, L., Muñoz, F., & Vargas, M. (2012). Managerial abilities: Evidence from religious mutual fund managers. *Journal of Business Ethics*, *105*(4), 503-517.

Filbeck, G., & Preece, D. (2003). Fortune's best 100 companies to work for in america: Do they work for shareholders? Journal of Business Finance & Accounting, 30(5-6), 771-797.

Filbeck, G., Gorman, R., & Zhao, X. (2013). Are the best of the best better than the rest? the effect of multiple rankings on company value. *Review of Quantitative Finance and Accounting, 41*(4), 695.

Freeman, R. (1984). Strategic Management: A Stakeholder Perspective (Piman, Boston).

Fulmer, I. S., Gerhart, B., & Scott, K. S. (2003). Are the 100 best better? an empirical investigation of the relationship between being a "great place to work" and firm performance. *Personnel Psychology, 56*(4), 965-993.

Galema, R., Plantinga, A., & Scholtens, B. (2008). The stocks at stake: Return and risk in socially responsible investment. *Journal of Banking & Finance, 32*(12), 2646-2654.

Gómez-Bezares, F., & Gómez-Bezares, F. R. (2012). Classic performance indexes revisited: axiomatic and applications. *Applied Economics Letters, 19*(5), 467-470.

Gómez-Bezares, F., Przychodzen, W., & Przychodzen, J. (2016). Corporate sustainability and shareholder wealth—Evidence from British companies and lessons from the crisis. *Sustainability*, *8*(3), 276.

Halbritter, G., & Dorfleitner, G. (2015). The wages of social responsibility—where are they? A critical review of ESG investing. Review of Financial Economics, 26, 25-35.

Henriksson, R., Livnat, J., Pfeifer, P., Stumpp, M., & Zeng, G. *ESG Literature Review.* Working paper

Herz, B., & Rogers, J. (2016). Measuring What Matters: Industry Specificity Helps Companies and Investors Gain Traction on Sustainability. *Journal of Applied Corporate Finance*, *28*(2), 34-38.

Hörisch, J., Ortas, E., Schaltegger, S., & Álvarez, I. (2015). Environmental effects of sustainability management tools: An empirical analysis of large companies. *Ecological Economics, 120*, 241-249.

Humphrey, J. E., Lee, D. D., & Shen, Y. (2012). Does it cost to be sustainable? Journal of Corporate Finance, 18(3), 626-639.

Javed, M., Rashid, M. A., & Hussain, G. (2016). When does it pay to be good–A contingency perspective on corporate social and financial performance: would it work?. *Journal of cleaner production*, *133*, 1062-1073.

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance*, *48*(1), 65-91.

Kempf, A., & Osthoff, P. (2007). The effect of socially responsible investing on portfolio performance. European Financial Management, 13(5), 908-922.

Khan, M., Serafeim, G., & Yoon, A. (2016). Corporate sustainability: First evidence on materiality. *The accounting review*, *91*(6), 1697-1724.

Lee, J., Yen, P., & Chan, K. C. (2013). Market states and disposition effect: Evidence from Taiwan mutual fund investors. Applied Economics, 45(10), 1331-1342.

Leite, P., & Cortez, M. C. (2015). Performance of European socially responsible funds during market crises: Evidence from France. *International Review of Financial Analysis*, *40*, 132-141.

Louche, C., & Lydenberg, S. (2006, July). Socially responsible investment: Differences between Europe and the United States. In Proceedings of the International Association for Business and Society (Vol. 17, pp. 112-117).

Loughran, T., & Ritter, J. R. (2000). Uniformly least powerful tests of market efficiency. *Journal of Financial Economics, 55*(3), 361-389.

Lu, W., Chau, K., Wang, H., & Pan, W. (2014). A decade's debate on the nexus between corporate social and corporate financial performance: A critical review of empirical studies 2002–2011. Journal of Cleaner Production, 79, 195-206.

Man, C. K. (2017). A Succinct Survey of Corporate Social Responsibility: Definition, Theory and Economic and Capital Market Repercussions. *Academy of Management Journal, Forthcoming*

Margolis, J. D., Elfenbein, H. A., & Walsh, J. P. (2009). Does it pay to be good... and does it matter? A meta-analysis of the relationship between corporate social and financial performance. *And does it matter*.

Mollet, J. C., & Ziegler, A. (2014). Socially responsible investing and stock performance: New empirical evidence for the US and European stock markets. Review of Financial Economics, 23(4), 208-216.

Mollet, J. C., von Arx, U., & Ilić, D. (2013). Strategic sustainability and financial performance: exploring abnormal returns. Journal of Business Economics, 83(6), 577-604.

MSCI. *2019 ESG Trends to Watch.* Available online: https://www.msci.com (accessed on 4 March 2019).

Neher, A. L., & Hebb, T. (2015). The responsible investment atlas–an introduction. *The Routledge Handbook of Responsible Investment*, 53-57. Routledge, New York.

Newey, Whitney K.; West, Kenneth D. A. Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation consistent Covariance Matrix, 1986.

Nofsinger, J., & Varma, A. (2014). Socially responsible funds and market crises. *Journal of Banking & Finance*, *48*, 180-193.

Orlitzky, M., Schmidt, F. L., & Rynes, S. L. (2003). Corporate social and financial performance: A meta-analysis. Organization Studies, 24(3), 403-441.

Osburg, T. (2013). Social innovation to drive corporate sustainability. In *Social Innovation* (pp. 13-22). Springer, Berlin, Heidelberg.

Pagan, A. R., & Sossounov, K. A. (2003). A simple framework for analysing bull and bear markets. Journal of Applied Econometrics, 18(1), 23-46.

Sandberg, J., Juravle, C., Hedesström, T. M., & Hamilton, I. (2009). The heterogeneity of socially responsible investment. *Journal of Business Ethics*, *87*(4), 519.

Scholtens, B. (2008). A note on the interaction between corporate social responsibility and financial performance. *Ecological economics*, *68*(1-2), 46-55.

Sheldon, O. (1924). The social responsibility of management. *The philosophy of management*, 70-99.

Ullmann, A. A. (1985). Data in search of a theory: A critical examination of the relationships among social performance, social disclosure, and economic performance of US firms. *Academy of management review*, *10*(3), 540-557.

USSIF: The Forum for Sustainable and Responsible Investment (2018)**.** *Report on US Sustainable, Responsible and Impact Investing Trends.* Available online: https://www.ussif.org (accessed on 18 February 2019).

Van de Velde, E., Vermeir, W., & Corten, F. (2005). Corporate social responsibility and financial performance. Corporate Governance: The international journal of business in society, 5(3), 129-138.

Wood, D. J., & Jones, R. E. (1995). Stakeholder mismatching: A theoretical problem in empirical research on corporate social performance. *The International Journal of Organizational Analysis*, *3*(3), 229-267.

Appendix A. Stock exchanges and associated countries

Table 6-4. Ap. A. Stock exchanges and associated countries

Stock Exchange	Country
VIE	AUSTRIA
BRU	BELGIUM
CPH	DENMARK
HEL	FINLAND
PAR	FRANCE
FRA	GERMANY
ATH	GREECE
ISE	ITALY
AMS	NETHERLANDS
WAR	POLAND
MAD	SPAIN
STO	SWEDEN
ZHR	SWITZERLAND
LON	UNITED KINGDOM
NAS	UNITED STATES
NYSE	UNITED STATES

Appendix B. Descriptive statistics

Table 6-5. Ap. B. Descriptive statistics

Panel A. EU market							
Materiality	Mean	Median	Maximum	Minimum	Std. Dev.	Skewness	Kurtosis
High [10]	0.0066	0.0076	0.1890	-0.2555	0.0596	-0.4564	5.2232
Low [10]	0.0023	-0.0031	0.2132	-0.1891	0.0600	0.0044	4.2202
High [20]	0.0066	0.0065	0.1990	-0.2826	0.0623	-0.6109	5.9463
Low [20]	0.0039	-0.0027	0.2281	-0.2245	0.0604	-0.0688	4.9102
High [30]	0.0059	0.0057	0.2144	-0.2676	0.0624	-0.4746	5.5288
Low [30]	0.0042	0.0000	0.2311	-0.2347	0.0621	-0.1291	4.9761
All Categories	Mean	Median	Maximum	Minimum	Std. Dev.	Skewness	Kurtosis
High [10]	0.0075	0.0057	0.2279	-0.2754	0.0616	-0.5735	6.3926
Low [10]	0.0045	0.0044	0.1971	-0.2279	0.0605	-0.2921	4.1340
High [20]	0.0068	0.0079	0.2295	-0.2618	0.0619	-0.4088	5.7437
Low [20]	0.0038	0.0036	0.2279	-0.2352	0.0611	-0.1865	4.9190
High [30]	0.0064	0.0066	0.2247	-0.2687	0.0622	-0.4808	5.7885
Low [30]	0.0042	0.0019	0.2284	-0.2463	0.0622	-0.2447	5.1205
Panel B. US market							
Materiality	Mean	Median	Maximum	Minimum	Std. Dev.	Skewness	Kurtosis
High [10]	0.0059	0.0126	0.1976	-0.2198	0.0594	-0.5443	4.5782
Low [10]	0.0040	0.0095	0.1990	-0.2494	0.0569	-0.5459	5.6762
High [20]	0.0052	0.0145	0.2139	-0.2218	0.0584	-0.4217	4.8545
Low [20]	0.0041	0.0084	0.1840	-0.2180	0.0541	-0.5105	4.9796
High [30]	0.0048	0.0120	0.2141	-0.2321	0.0580	-0.4654	5.2047
Low [30]	0.0046	0.0089	0.1956	-0.2217	0.0543	-0.5031	5.3815
All Categories	Mean	Median	Maximum	Minimum	Std. Dev.	Skewness	Kurtosis
High [10]	0.0058	0.0114	0.1959	-0.2388	0.0577	-0.5539	5.2144
Low [10]	0.0057	0.0095	0.2209	-0.2235	0.0551	-0.3189	5.7652
High [20]	0.0055	0.0125	0.2074	-0.2432	0.0573	-0.5475	5.5786
Low [20]	0.0049	0.0084	0.2123	-0.2281	0.0552	-0.4382	5.7258
High [30]	0.0056	0.0124	0.2087	-0.2368	0.0562	-0.5168	5.6779
Low [30]	0.0051	0.0090	0.2096	-0.2233	0.0546	-0.4314	5.6101

General conclusions and further research

This section includes a summary of the main findings and conclusions of this Doctoral Thesis as well as some further research lines.

In Chapter 1 we test whether it is possible to consistently achieve extra-financial returns by means of a sector strategy using the Fama and French model (1992) as a basis for decision-making. Our results show that this strategy has a limited utility since it is not possible to achieve positive extra-financial performance systematically.

In Chapter 2 we test the effectiveness of the Fama and French model (1992) by taking month-to-month data and reforming the value and size portfolios at the end of each month, aiming to develop a more dynamic and adaptable tool. Our results show that the Fama and French model (1992) can be converted into a more flexible and dynamic tool since the construction of the risk factors taking monthly data allows us to adapt them more recurrently, resulting in an improved ability to capture the variations that may arise in the characteristics of firms.

In Chapter 3 we assess the performance of SRI portfolios formed on the basis of the Global 100 list over the period 2005 to 2014. Previous retail investor research is focused on specific countries but we provide evidence of SRI financial performance at the worldwide level as well as at the regional level, for 5 regions (North America, Europe except UK, United Kingdom, Pacific region and Emerging markets). Additionally, as recent research shows that SRI performance can differ across market states, we analyse SRI portfolio performance in periods of bull and bear markets. Our results show that the Global-100 portfolio outperforms conventional investments. The results on SRI financial performance at the regional level show statistical differences in the financial performance of the SRI regional portfolios. The regional analysis allows us to conclude that the performance of the Global-100 portfolio is mostly influenced by three specific regional portfolios: North America and Europe ex-UK (positive impact) and emerging markets (negative impact). As to the differences in performance between SRI and conventional investments across different market states, the results show that the financial performance in bear market periods is neutral for both portfolios. However, the Global-100 portfolio outperforms the S&P 100 Index in up markets. In sum, our empirical evidence indicates that socially conscious retail investors are able to implement a SRI strategy that outperforms conventional investments. In addition, the different results uncovered at the regional level suggest that country-specific factors may affect the relationship between corporate social and financial performance.

In Chapter 4 we investigate the financial performance of international stock portfolios based on CSR criteria. Using an international dataset of companies between 2002 and 2017, we extend the evidence on SRI portfolio performance to North America, Europe, Japan, and Asia Pacific. Our results show that, in most cases, there are no statistical significant differences in the performance of portfolios of firms with higher sustainability scores and those with lower sustainability scores. Nevertheless, we find particular influential effects of screening processes within regions. In terms of the Governance dimension, we observe contrasting effects: some EU high-rated portfolios underperform low-rated ones, while some AP high-rated portfolios outperform low-rated ones. On the Social dimension, we find a strong positive effect in EU firms: high-ranked firms outperform their low-ranked counterparts whatever the cut-off used. Screening processes based on the Environment dimension do not uncover significant performance differences. In addition, we analyse how SRI portfolios perform across different market states. While for the full sample period we found just a few cases with differences in performance between high- and low-rated portfolios, assessing the financial performance over different market phases uncovers several portfolio performance differentials. The differences are especially notable in EU over bull markets. Our findings suggest that significant performance differences among SRI portfolios of different regions is a result of firms of different regions being influenced by distinct economic conditions. In sum, these results suggest that the impact of social screening in portfolio performance is market state and geographically dependent.

In Chapter 5 we evaluate the financial performance of government bond portfolios formed according to ESG criteria. Using the RobecoSAM information to classify the government bonds according to ESG performance, we assess financial differences between high- and low-ranked government bonds over the period June 2006 to December 2017. Our results show that high-rated portfolios outperform low-rated ones under any SRI demanding level (cut-off), although differences are not significant. These findings are in line with most previous studies which find that SRI perform similarly to conventional investments. The absence of significant differences has been widely considered as a relevant finding. All in all, our evidence indicates that an investor can satisfy ESG concerns without sacrificing financial performance by investing in government bonds. In this regard, given that SRI investor claims have led firms to be more concerned with their corporate social responsibility strategies (Ullmann, 1985), investors screening government bonds according to their sustainability scores could

influence countries in terms of ESG guiding principles. Our results suggest that SRI can be used as a tool to enhance ESG policy of countries.

In Chapter 6 we assess the financial performance of stock portfolios formed according to material and immaterial CSR issues. Our dataset includes companies from North America and Europe. Evaluating firms from North America and Europe is particularly interesting given the heterogeneity in the patterns of development of SRI across countries (Neher and Hebb, 2015). Our results in the EU market show that materiality is relevant to find the best and worst firms both in terms of ESG and of financial performance whereas we do not find a significant positive effect of materiality on the US firm's financial performance. Despite results from the US market, we consider that materiality is relevant to both firms and investors. First, it allows firms to focus their sustainability strategies on the most important issues. Second, it permits investors to evaluate portfolio exposure to specific material and immaterial sustainability risks and opportunities. However, the material entity-specific information seems to be still imprecise because, in the US market, material scores do not enable us to discriminate the best firms both in terms of ESG and of financial performance.

We consider that further research would be worthwhile to enlarge on the knowledge in these fields. As for chapters 1 and 2, it would be of interest to evaluate the validity of multi-factor models as financial performance measures. Many researchers have accepted these measures as suitable ones but as Barber and Lyon (1997), Fama (1998), and Loughran and Ritter (2000) note, the magnitude and sometimes even the sign of the long-run abnormal returns are sensitive to alternative measurement methodologies.

With respect to chapter 3, it would be interesting to enlarge the evidence on the possibilities of socially responsible retail investors to perform 'good' in terms of ESG and well in terms of financial performance. According to the most recent report of the GSIA (2018), the proportion of retail investors rose from 20% to 25% during the period 2016 to 2018. This shows that the tendency noted in chapter 3 for the period 2014 to 2016 continues. In this regard, it would be of particular interest to supply information sources - list of stocks (e.g., the Global-100 list) - which are freely available to the public and any retail investor can easily access them.

As for chapter 4, it would be of interest to go on assessing the financial performance of SRI synthetic portfolios in an international context. Considering the growth of socially responsible investments in international capital markets and the intensifying global competition, the valuation implications of sustainability in an international context is of

practical interest to management, investors and regulators worldwide. In this evaluation, since previous studies that address the performance of socially screened synthetic portfolios suffer from some limitations and inconsistencies, future studies should pay attention to the seven shortcomings identified in chapter 4.

With respect to chapter 5, it would be interesting to evaluate the particular effect of each ESG dimension on the financial performance of SRI government bond portfolios. Moreover, different maturities of government bonds could be considered to form portfolios, as well as including more countries, especially developing countries, and evaluating a longer sample period. In addition, since previous literature on SRI investment funds and SRI stock portfolios finds that different market states (e.g., expansion and recession) affect the financial performance of SRI, researchers could evaluate this concern about SRI government bond portfolios. Specific SRI issues assessed previously on other assets could actually be analysed from now in this context.

In chapter 6, it would be of interest to extend the analysis to a longer period in order to assess whether it affects the results. In addition, we consider that material issues should still be defined in an appropriate manner so that investors and firms can use them as an appropriate tool.

Conclusiones generales y líneas de investigación futuras

Este apartado incluye un resumen de los principales resultados y conclusiones de esta Tesis Doctoral, así como algunas líneas de investigación adicionales.

En el Capítulo 1 se comprueba si es posible obtener de forma consistente rendimientos extraordinarios mediante una estrategia sectorial basada en el modelo de Fama y French (1992) como base para la toma de decisiones. Nuestros resultados muestran que esta estrategia tiene una utilidad limitada, ya que no es posible alcanzar un rendimiento extraordinario positivo de forma sistemática.

En el Capítulo 2 probamos la eficacia del modelo Fama y French (1992) tomando datos mensuales y reformando las carteras de valor y tamaño al final de cada mes, con el objetivo de desarrollar una herramienta más dinámica y adaptable. Nuestros resultados muestran que el modelo pueden convertirse en una herramienta más flexible y dinámica ya que la construcción de los factores de riesgo a partir de los datos mensuales nos permite adaptarlos de forma más recurrente, lo que se traduce en una mayor capacidad para captar las variaciones que pueden surgir en las características de las empresas.

En el Capítulo 3 evaluamos el rendimiento de carteras ISR formadas sobre la base de la lista Global-100 durante el período 2005-2014. Las investigaciones previas sobre los inversores minoristas se centra en países específicos sin embargo nosotros proporcionamos evidencia tanto a nivel internacional como regional (Norteamérica, Europa excepto Reino Unido, Reino Unido, la región del Pacífico y los mercados emergentes). Además, dado que las investigaciones recientes muestran que el rendimiento de la ISR puede diferir de un estado del mercado a otro, analizamos el rendimiento de las carteras ISR en períodos de mercados alcistas y bajistas. Nuestros resultados muestran que la cartera de Global-100 supera a las inversiones convencionales. Los resultados sobre el desempeño financiero de la ISR a nivel regional muestran diferencias estadísticas para carteras regionales. El análisis regional nos permite concluir que el desempeño de la cartera de Global-100 está influenciado principalmente por tres carteras regionales específicas: América del Norte y Europa ex-Reino Unido (impacto positivo) y mercados emergentes (impacto negativo). En cuanto a las diferencias de rendimiento entre la ISR y las inversiones convencionales en los distintos estados del mercado, los resultados muestran que el rendimiento financiero en períodos de mercado bajista es neutro para ambas carteras, sin embargo, la cartera de Global-100 supera al índice S&P 100 en los mercados al alza. En resumen, nuestra evidencia empírica indica que los inversores minoristas con conciencia social son

capaces de implementar una estrategia de ISR que supera a las inversiones convencionales. Además, los diferentes resultados descubiertos a nivel regional sugieren que los factores específicos de cada país pueden afectar la relación entre el desempeño social y financiero de las empresas.

En el Capítulo 4 investigamos el desempeño financiero de carteras de acciones basadas en criterios de RSC a nivel internacional. Utilizando un conjunto de datos de empresas entre 2002 y 2017 ampliando la evidencia a Norteamérica, Europa, Japón y Asia-Pacífico. Nuestros resultados muestran que, en la mayoría de los casos, no existen diferencias estadísticamente significativas en el desempeño de las carteras de las empresas con mejores puntuaciones de sostenibilidad y de aquellas con puntuaciones de sostenibilidad más bajas. Sin embargo, sí que encontramos efectos particularmente influyentes de los procesos de selección dentro de las regiones. En cuanto a la dimensión de Gobernanza, observamos efectos variados: algunas carteras de alta calificación de la UE tienen un rendimiento inferior a las de baja calificación, mientras que algunas carteras de alta calificación de AP tienen un rendimiento superior a las de baja calificación. En cuanto a la dimensión social, encontramos un fuerte efecto positivo en las empresas de la UE: las empresas de alto nivel superan a sus homólogas de bajo nivel, sea cual sea el corte utilizado. Los procesos de cribado basados en la dimensión Medioambiente no revelan diferencias de rendimiento significativas. Además, analizamos cómo se comportan las carteras de ISR en los diferentes estados del mercado. Mientras que durante todo el período de muestra encontramos sólo unos pocos casos con diferencias de rendimiento entre carteras de alta y baja calificación, la evaluación del rendimiento financiero en diferentes fases del mercado revela varias diferencias de rendimiento. Las diferencias son especialmente notables en los mercados alcistas de la UE. Nuestros hallazgos sugieren que las diferencias de rendimiento entre las carteras ISR de las distintas regiones se deben a que las empresas de las distintas regiones se ven influenciadas por condiciones económicas distintas. En resumen, estos resultados sugieren que el impacto social en el rendimiento de la cartera es diferente en cada región y para cada estado del mercado.

En el Capítulo 5 evaluamos el rendimiento financiero de las carteras de deuda pública formadas de acuerdo con criterios ESG. Utilizando información de RobecoSAM para clasificar los bonos del Estado según el rendimiento ESG, evaluamos las diferencias financieras entre bonos gubernamentales de alta y baja calificación durante el período comprendido entre junio de 2006 y diciembre de 2017. Nuestros resultados muestran

que las carteras de alta calificación superan a las de baja calificación bajo cualquier nivel de exigencia (cut-off) de la ISR, aunque las diferencias no son significativas. Estos resultados coinciden con los de la mayoría de los estudios anteriores, en los que se ha comprobado que la ISR funciona de forma similar a las inversiones convencionales. La ausencia de diferencias significativas ha sido ampliamente considerada como un hallazgo relevante. En general, nuestra evidencia indica que un inversor puede satisfacer sus preocupaciones ESG sin sacrificar el rendimiento financiero invirtiendo en bonos gubernamentales. A este respecto, dado que los intereses de los inversores en ISR han llevado a las empresas a preocuparse más por sus estrategias de responsabilidad social corporativa (Ullmann, 1985), los inversores que seleccionan bonos gubernamentales en función de sus puntuaciones de sostenibilidad podrían influir en los países en términos de sus principios rectores con respecto a la ESG. Nuestros resultados sugieren que la ISR puede utilizarse como una herramienta para mejorar la política ESG de los países.

En el Capítulo 6 se evalúa el rendimiento financiero de carteras de valores formadas en función de cuestiones de RSC materiales e inmateriales. Nuestro conjunto de datos incluye empresas estadounidenses y europeas. La evaluación de empresas estadounidenses y europeas es particularmente interesante dada la heterogeneidad en las pautas de desarrollo de la ISR en los distintos países (Neher y Hebb, 2015). Nuestros resultados en el mercado de la UE muestran que la materialidad es relevante para encontrar las mejores y las peores empresas, tanto en términos de ESG como de rendimiento financiero, mientras que no encontramos un efecto positivo significativo de la materialidad en el rendimiento financiero de las empresas estadounidenses. A pesar de los resultados del mercado estadounidense, consideramos que la materialidad es relevante tanto para las empresas como para los inversores. En primer lugar, permite a las empresas centrar sus estrategias de sostenibilidad en las cuestiones más importantes y en segundo lugar, permite a los inversores evaluar la exposición de sus carteras a riesgos y oportunidades de sostenibilidad materiales e inmateriales específicos. Sin embargo, la información material específica de cada empresa parece ser todavía imprecisa porque en el mercado estadounidense las puntuaciones materiales no nos permiten discriminar las mejores empresas tanto en términos de ESG como de rendimiento financiero.

Por último, consideramos que sería interesante seguir investigando en la línea de esta Tesis Doctoral para ampliar el conocimiento por varios motivos. En cuanto a los capítulos 1 y 2, sería interesante evaluar la validez de los modelos multifactoriales como

medidas de performance financiera. Muchos investigadores han aceptado estas medidas como adecuadas, pero como Barber y Lyon (1997), Fama (1998), y Loughran y Ritter (2000) notan, la magnitud y a veces incluso el signo de los rendimientos anormales a largo plazo son sensibles a metodologías de medición alternativas.

Por lo que se refiere al capítulo 3, sería interesante ampliar las pruebas sobre las posibilidades de que los inversores minoristas socialmente responsables obtengan resultados buenos en términos de ESG y buenos en términos de resultados financieros. Según el informe más reciente de la GSIA (2018), la proporción de inversores particulares aumentó del 20% al 25% durante el período 2016-2018. Esto demuestra que la tendencia observada en el capítulo 3 para el período 2014-2016 continúa. A este respecto, sería de particular interés proporcionar fuentes de información -listas de acciones (como por ejemplo la lista Global-100)- que estén a disposición del público de forma gratuita y que cualquier inversor minorista pueda acceder fácilmente a ellas.

En cuanto al capítulo 4, sería interesante seguir evaluando los resultados financieros de las carteras sintéticas de ISR en un contexto internacional. Teniendo en cuenta el crecimiento de las inversiones socialmente responsables en los mercados internacionales de capitales y la intensificación de la competencia mundial, las implicaciones de la sostenibilidad en un contexto internacional son de interés práctico para las empresas, los inversores y los reguladores de todo el mundo. En esta evaluación, dado que los estudios previos que abordan el desempeño de las carteras sintéticas sometidas a cribado social adolecen de algunas limitaciones e incoherencias, los estudios futuros deberían prestar atención a las deficiencias identificadas en el capítulo 4.

Con respecto al capítulo 5, sería interesante evaluar el efecto particular de cada dimensión de la ESG en el rendimiento financiero de las carteras de deuda pública. Asimismo, se podrían considerar diferentes vencimientos de los bonos gubernamentales para formar las carteras, además de incluir a más países, especialmente a los países en desarrollo, así como evaluar un período de muestreo más largo. Dado que la literatura previa sobre fondos de inversión ISR y carteras de acciones ISR encuentra que los diferentes estados del mercado (por ejemplo, expansión y recesión) afectan el desempeño financiero de la ISR, los investigadores podrían evaluar esta preocupación sobre las carteras ISR de bonos gubernamentales. Las cuestiones específicas de ISR evaluadas anteriormente sobre otros activos podrían analizarse a partir de ahora en este contexto.

En el capítulo 6, sería interesante ampliar el análisis a un período más largo para evaluar si los resultados se ven afectados. Además, consideramos que los aspectos materiales deberían definirse de manera más adecuada para que los inversores y las empresas puedan utilizarlos como una herramienta de decisión.