Hidden Gem or Fool's Gold: Can Passive ESG ETFs Outperform the Benchmarks?

Ariadna Dumitrescu*[a], Jesse Järvinen[b], Mohammed Zakriya[c]

[a]*ESADE Business School, Universitat Ramon Llull*
Av. Pedralbes, 60-62, 08034 Barcelona, Spain

[b]*IESEG School of Management,*
3 rue de la Digue, 59000 Lille, France

[c]*IESEG School of Management, Univ. Lille*
CNRS, UMR 9221 - LEM - Lille Économie Management,
3 rue de la Digue, 59000 Lille, France

This Version: December 2021

Abstract

Using a unique and extensive dataset of 121 socially responsible investing (SRI) equity exchange-traded funds (ETFs) from January 2010 to December 2020, this paper studies how passive SRI ETFs perform in comparison to their non-SRI benchmarks. Over the full sample period, our results show that an equally weighted SRI ETF portfolio underperforms its benchmark portfolio. Further, on examining how the two portfolio performs in the second half of our sample period, we find no significant difference between them. Moreover, in the last two years, we find that the SRI ETF portfolio significantly outperforms the benchmark. We also show that positive screening (or, inclusion) rather than negative screening (exclusion) is the ETF investment strategy that can beat the benchmark portfolio. In particular, the environmental inclusion screens provide significantly large abnormal returns. Lastly, we find that increasing competition and declining market concentration is important for SRI ETFs' performance.

Keywords: ESG, ETFs, SRI, socially responsible investments, sustainability

JEL Classification: G32, G34, L21, M14, Q56.

*Corresponding author, Tel.: +34 934952191
Email addresses: ariadna.dumitrescu@esade.edu (Ariadna Dumitrescu*), jesse.jarvinen@ieseg.fr (Jesse Järvinen), m.zakriya@ieseg.fr (Mohammed Zakriya)

"Talking about ETFs is like talking about people. There are good ones, and there are bad ones."

–JOHN C. BOGLE

1. Introduction

Socially responsible investing (SRI) or environmental, social, and governance (ESG) investing has grown substantially during recent years in asset size with more and more investment products being frequently re-oriented to promote the idea of doing good while doing well. At the same time, as the popularity of passive investing, and more specifically, the exchange-traded funds (ETFs) grew (Elton et al., 2019), they have adapted their assets under management and product offerings to offer SRI-oriented alternative investments in the form of passive SRI ETFs. These ETFs are important investment options for investors to consider because they offer a transparent and cost-efficient way of implementing SRI in practice.[1] In other words, passive SRI ETFs encapsulate together the twin cost advantages arising from the lower expenses of passive funds versus the active funds (French, 2008) and from the cost savings due to ETFs' lower management fees when compared to traditional mutual funds (Lettau and Madhavan, 2018).[2] Given that these factors have lead to a substantive growth in the popularity of passive ETFs, this paper focuses on evaluating the performance of passive SRI ETFs in comparison to their passive non-SRI benchmarks, while also examining factors that act as drivers of their value creation or destruction.

This paper contributes to the growing debate on the ability (or inability) of SRIs to outperform the markets (Gil-Bazo et al., 2010; Nofsinger and Varma, 2014; Riedl and Smeets, 2017; Omura et al., 2021). The novelty comes from the fact that we specifically study passive equity ETFs as an investment vehicle so that the benefits (or, costs) of sustainable investing or SRI are cleanly captured by essentially leaving active management

[1]For the U.S. funds, the 2019 Morningstar report shows that while the fund fees for both active and passive funds have been on a decline over the years, the passive funds still continue to be 0.5% cheaper than their active counterparts.

[2]Not that we use SRI ETFs and ESG ETFs interchangeably throughout the paper to represent the group of equity ETFs that implement SRI- or ESG-oriented investment strategies.

skills out of the purview. Fundamentally, as Climent and Soriano (2011) note, "[i]t might be the case that some SRI funds owe their success/failure more to standard investment techniques (smaller firms, higher beta coefficients, ...) than to a particular ethical screening technique." Thus, by focusing on passive equity ETFs that usually merely replicate the holdings of existing indices, we aim to estimate the benefits of SRI funds in comparison to their non-SRI benchmarks by keeping it free of any active management-specific biases. Our analysis enables us to extend the existing literature in several ways. First, we provide an extensive outlook of the growing trends in SRI ETFs both in terms of their numbers and size using the longest and most comprehensive SRI ETFs sample so far. To the best of our knowledge, this is the first paper to track the performance of a substantially large set of passive equity SRI ETFs over a long period of time. Given the fact that SRI ETFs are relatively new financial instruments, these descriptives by themselves are informative for both investors and future researchers. Second, by analyzing the performance of passive SRI ETFs and their benchmarks, we are able to identify any time-specific trends observed in their performance. As passive SRI ETFs provide a relatively bias-free representation of SRI funds in general, this helps us understand whether SRI is "smart money" or not. Thus, we provide insights on a bigger question that SRI industry has to often tackle: should we simply treat SRI as an investment choice that serves investors' altruistic needs or do they also have the propensity to generate abnormal returns? Third, for each SRI ETF in our sample, we manually collect their SRI implementation strategies (or screening processes) and shed light on specific SRI strategies that can potentially outperform the benchmarks.

We provide novel insights on SRI performance by highlighting how SRI-oriented investors could benefit from passive asset management while also maintaining social responsibility as their core investment philosophy. SRI ETFs are instruments that allow investors to passively implement sustainability-focused investment style. While there is much existing literature studying a combination of social responsibility and mutual funds, there is not much evidence on SRI ETFs' performance. As mentioned above, the growing demand for SRI products among investors means that passive ETFs have

emerged as a natural cost-effective solution for SRI-oriented investors. Thus, it is crucial to understand whether they can offer a better alternative to mutual funds or other similar investment products. In doing so, we also contribute to the broader literature that studies SRI-oriented mutual funds and investment. When it comes to passive SRI funds, literature has mainly focused on mutual funds and finds no consistent evidence for under- or over-performance in comparison to their benchmarks (e.g., Bauer et al., 2005; Renneboog et al., 2008b; Nofsinger and Varma, 2014). We extend this literature by studying a relatively newer category of SRI-based investment vehicle, and find that it does not consistently lose to its benchmark. In fact, the SRI ETFs show a potential to generate better alphas depending on the time period analyzed.

We identify a unique and comprehensive dataset of 121 U.S. based passive SRI ETFs and trace their performance over an 11-year period of January 2010 to December 2020 to compare how they fare over passive ETF counterparts that follow the S&P500 Index. Since our sample period covers the equity market crash accompanying the Covid-19 crisis, we are also indirectly providing insights on the differences in the performance of passive SRI ETFs during crisis and non-crisis years. Thus, we are able to empirically test whether SRI ETFs have any advantages over the non-SRI ETFs during the period of economic downturn by mitigating their downside risks (Nofsinger and Varma, 2014). This ESG-related resilience has been clearly established from firms' perspective (Dumitrescu and Zakriya, 2021), especially during the Covid-19 crisis (Albuquerque et al., 2020; Ding et al., 2021). However, whether similar resilience and CSR-related risk mitigation exists for the investors is yet to be firmly established. Since even the passive SRI ETFs may target several different types of stocks or indices depending on their focal objectives. For instance, while some ETFs may choose to distance themselves from certain "sin" stocks or industries, others might choose to ascertain their picks by considering ESG ratings, controversies, or standards. Furthermore, SRI ETFs may employ positive screens (inclusion criteria to identify good ESG firms) or negative screens (exclusion of ESG laggards). Thus, we examine different SRI ETFs by grouping them based on their ESG-related objectives and screens to identify if any specific ESG-based strategy can outperform passive

non-SRI ETFs. In other words, we also empirically tests which of the ESG attributes and strategies employed by the passive SRI ETFs can create value for investors.

The popularity and demand of U.S. equity SRI ETFs can be adjudicated from the fact that the number of SRI ETFs grows more than four times from the beginning of our sample period in 2010 to 2020. When it comes to the performance comparison, our analysis reveals that SRI ETFs do not always underperform their passive non-SRI benchmark ETFs. Instead, during the second half of the sample period (i.e., between 2015 and 2020) when this industry develops and flourishes, there is no statistically significant difference between their performance and that of the benchmarks. More importantly, we find that investors can potentially gain significant value from the SRI ETFs in the last two years of our sample. We posit that the growing popularity of the SRI ETFs and their higher resilience during the Covid-19 crisis may have played an important role in the SRI ETF portfolio's outperformance in recent years over its benchmark. This means that for investors considering SRI, ETFs, and passive asset management, individually and as a combined investment instrument, our results have important value implications. Furthermore, we find that only the SRI ETFs that incorporate Inclusion (positive screening) — and in particular, those that employ the Environmental Inclusion strategy — are drivers of abnormal returns. Given that management fees plays an important role in mutual funds performance (Gil-Bazo and Ruiz-Verdú, 2009), especially for specialized funds like SRI (Gil-Bazo et al., 2010), we control for the effect of fees on SRI ETFs' returns and find that all our main results remain unaffected.

The outperformance of SRI ETFs over their benchmarks documented in recent years could be consistent either with a) a demand-side explanation, i.e., as more investors become aware of and demand sustainability- or SRI-oriented assets, the SRI ETFs may be pressured to perform better, or b) a supply-side explanation, i.e., an increased market competition due to new entrants in recent years may have positively influenced the SRI ETF industry by improving the market quality. Our evidence is more consistent with the supply-side effect than that of the demand-side explanation. In particular, we find that a decline in SRI ETFs' market concentration (or, increase in their competition) is

associated with higher returns. On the other hand, an increase in sustainability rating for SRI ETFs, that generally is associated with an increased demand and fund inflows (Hartzmark and Sussman, 2019), has no effect on returns.

The rest of this study is structured as follows. Section 2 reviews the theoretical background including the differences between active and passive asset management, SRI, ETFs, and their investment strategies. This section also provides a quick summary of historical evolution of SRI ETFs, their investment criteria, and why their performances may differ from non-SRI ETFs. We also motivate our hypotheses alongside. Section 3, then, introduces the dataset used in this study, its collection from various sources, and the methodology and empirical models employed. Sections 4 and 5 present the results from our empirical analyses comparing the SRI ETFs and their passive ETF benchmarks. Section 6 examines possible explanations for SRI ETFs' returns. Lastly, Section 7 summarizes all the main findings and concludes.

2. ETFs and SRI ETFs

The fundamental function of investment is to obtain additional profit or income in the future, whether it is through passive or active strategies (Fabozzi and Markowitz, 2011). Passive asset management can potentially provide more value than active asset management given its lower fees and expense ratios (Arnott and Darnell, 2003). Investors generally prefer lower fees since higher fees destroy the expected returns (Gruber, 1996) given that expense ratios, portfolio turnover, and load fees affect the portfolio performance significantly negatively (Carhart, 1997). In case of active asset management, it is difficult to identify whether asset managers' characteristics rather than that of the managed funds drive their performances (Chevalier and Ellison, 1999; Kacperczyk et al., 2016). Thus, by focusing on passive investment funds, we aim at isolating the benefits of fund characteristic or objective by completely removing managerial stock picking or market timing abilities out of the equation.

Passive asset management and index investing have been grounding drivers of a relatively new investment innovation: The ETFs. Most ETFs typically seek to follow a specific index's performance, as an index mutual fund does, but ETFs vary in many cru-

cial ways from mutual funds. As French (2008) suggests, active asset management costs (using mutual funds) are high, and this steers investors towards passive asset management strategies that involve fewer transactions, hence, resulting in greater cost-efficiency. Thus, ETFs are a convenient and cost-effective way for small investors to reach special markets that would be too expensive or otherwise complicated to access (Delfeld, 2007; Lettau and Madhavan, 2018). Additionally, the risks of mistimed investment or of mistakenly picking wrong stocks in actively managed portfolios can be avoided. Note that although the SRI ETFs usually have higher sustainability rankings than the conventional counterparts, they could also have comparatively higher costs due to stricter screening strategies. ETFs have recently shown remarkable year-on-year growth in assets, diversity, and market significance to now become an important financial asset, and investors are well aware of this (Elton et al., 2019). Thus, ETFs' popularity increased amongst the investors (Lettau and Madhavan, 2018) despite the debate surrounding their persistent short-termism (Barber and Odean, 2000; Dellva, 2001).

2.1. SRI: Opportunities and Challenges

It has not been long since the investment style of socially responsible investing (SRI) started developing and going mainstream. More recently, the asset managers have increasingly began accepting SRI practices and incorporating ESG issues into their investment decision-making process and analysis (Riedl and Smeets, 2017). SRI funds have grown substantially during recent years both in terms of the assets under management and the number of available products that incorporate the idea of "doing good while doing well." Socially responsible funds began by initially excluding stocks that deal with tobacco, alcohol, gambling, and fossil resources, usually referred to as "sin stocks." Further, the SRI funds use different screening strategies like inclusion (positive screening), where the asset manager selects stocks with the best ESG attributes, and exclusion (negative screening), where the asset manager rules out the stocks with the worst ESG attributes (Nofsinger and Varma, 2014).

In general, ESG issues can be important for long-term survival and value creation of firms (Fatemi et al., 2015). Much of the SRI landscape (that includes alternative nomen-

clatures such as sustainable investing, ESG investing, or impact investing) generally aim at benefiting from such long-run prospects of ESG. More specifically, SRI seeks to deliver returns without compromising on the long-term impact of companies' business policies on society and on the environment. So, SRI investors are not only seeking financial returns, but they are also targeting environmental and social outcomes. When interest and awareness towards these issues rise among investors and there is increasing demand from legislators, there are complementing pressures on companies to integrate socially responsible and sustainable matters into their business processes (Renneboog et al., 2008b). Essentially, this self-reinforcing mechanism from SRI-oriented investors to their portfolio companies is one of the important factors driving the growing demand for SRI funds (Joliet and Titova, 2018). No matter what their effect is on companies, previous evidence suggests that environmental, social, and governance issues impact companies' value, and managers can no longer ignore this (Waddock et al., 2002).

From investors' perspective, ESG characteristics are becoming increasingly relevant to investment decisions because SRI investors have to choose from a limited investment opportunity set (Benson and Humphrey, 2008). Despite the fact that these investors themselves may be motivated by their intrinsic social preferences and social signalling needs (Riedl and Smeets, 2017), their possibly profit-driven hidden quest for smart money effect cannot be ignored (Renneboog et al., 2008a). Thus, SRI investors are also increasingly demanding more and more cost-efficient, liquid, and easily distributed investment solutions that parallelly integrate SRI principles.

SRI encompasses a continuum of new financial instruments targeted at issues related to sustainability like climate change, pollution, environmental sustainability, ethical awareness, and matters like consuming resources more effectively (for a discussion of history and evolution of SRI, see Schueth, 2003; Renneboog et al., 2008b).[3] In the U.S., the SRI market has grown from managing 639 billion US$ assets in 1995 to now totalling

[3]SRI is an umbrella term representing an investment decision-making process that merges social, environmental, and ethical consideration along with investment returns (Renneboog et al., 2008b; Brzeszczyński and McIntosh, 2014). In this regard, sustainable investing, ESG investing, impact investing, ethical investing, or green investing, all tend to be used interchangeably to represent a broadly similar investment strategy.

over 17.1 trillion US$ as shown in the U.S. SIF Foundation's 2020 biennial Report on US Sustainable, Responsible, and Impact Investing Trends.

Why should investors care about ESG or corporate social responsibility? The traditional finance theory considers that companies' objective is to maximize their shareholders' value and, hence, investors' returns (Bodie et al., 2014). This is consistent with Friedman (1970) arguments that focusing on other stakeholders' needs may result in compromising shareholders' wealth. Notwithstanding these arguments, many studies have consistently shown the potential of corporate social responsibility and stakeholder orientation in creating value for investors (Kempf and Osthoff, 2007; Gil-Bazo et al., 2010; Dumitrescu and Zakriya, 2021). Alongside increasing investors' financial value, CSR and sustainability can also create non-financial value by satisficing investors especially when they have non-pecuniary ESG preferences (Bollen, 2007). In fact, some investors may even favor ESG and sustainability preferences, despite possibilities of suffering financial losses (Rivoli, 1995; Renneboog et al., 2011).

One of the biggest challenges for SRI-oriented investors or fund managers is whether to integrate all of the ESG factors into the investment decision-making process or to focus on a specific ESG dimension. In general, the SRI investing and portfolio construction process considers risk and returns patterns as any other investment strategy but also ensures that the ESG factors are included for evaluating the portfolio performance. To do so, some investors may seek to invest only in companies with superior ESG policies and include them in their portfolios (positive or best-in-class screening), while other may exclude companies with poor ESG record or those that frequently report ESG controversies (negative or exclusionary screening). Many socially responsible funds avoid investing in industries like tobacco, weapons, alcohol, and gambling.[4] While negative screens are the oldest form of SRI strategy, positive screening has grown in popularity lately. This is mainly due to different ESG and sustainability scores being made available to investors

[4]A policy that restricts "sin-stocks" is a form of negative screening. It can also be referred to as product-based screening. For example, such firms may be involved in selling weapons, tobacco, or alcohol. These stocks are known to yield premium returns to compensate for their reputation risks (Hong and Kacperczyk, 2009), but recent findings have shown that sin-stock need not necessarily always yield abnormal returns (Blitz and Fabozzi, 2017).

to help them select superior ESG assets.

The most popular limitation or critique of SRI is that SRI funds may incur diversification costs due to limited possibilities for optimal investment allocation (Guenster, 2012; Girard et al., 2007) as the ESG screens constrain the investment opportunity set. Thus, SRI investors may carry more unsystematic risk because of suboptimal diversification (Barber and Odean, 2000), especially when positive alpha firms are excluded through ESG screening (à la Hong and Kacperczyk, 2009). However, this is not always the case as SRI funds have also been shown to not differ significantly from conventional funds in terms of diversification, asset allocation, and portfolio holdings (Bello, 2005). When it comes to performance as well, SRI funds may not always be compromising on stock returns to integrate ESG in the investment decisions. For instance, investing using strict SRI screens might help investors pick stocks that are better performers because a specific ESG dimension or screen may create value. Past literature has demonstrated that this is indeed the case. ESG-based investments have shown potential to generate abnormal returns, for instance, in case of environmentally responsible firms (Derwall et al., 2005; Kempf and Osthoff, 2007), firms with high employee satisfaction (Edmans, 2011; Derwall et al., 2011), and firms with good corporate governance (Bebchuk et al., 2009).

2.2. SRI ETFs: Past, Present, and Future

Although the first ETF, "The Spider", was launched in 1993, it was only a decade later when the first SRI ETF came around. The MSCI (KLD) USA ESG Select NR USD was launched on January 28, 2005. Meziani (2014) explains that this slow start of SRI-based financial innovations coincided with the overall slow start of ETFs themselves. Once investors began adopting ETFs, it was only a matter of time that the SRI-based strategy was integrated into ETFs. SRI ETFs are investment instruments that invest primarily in socially responsible assets, defined as stocks, commodities, or fields of industries that exhibit positive ESG characteristics. Thus, they hold a collection of socially responsible companies in their portfolio (Chakrabarty et al., 2017).

Very few studies have examined the potential for value creation and financial outperformance from SRI or ESG ETFs. All of them are severely constrained either by

limited or short time series of data or by small sample sizes due to the emerging nature of these financially innovative products. Sabbaghi (2011) examines their emergence from their debut in January 2005 through October 2009 by studying 15 SRI ETFs, and shows that while median returns for these ETFs are positive, they are not immune to extreme downside risks like that during the 2008 financial crisis. Meziani (2014) measures the ESG ETFs' potential to add value relative to the more traditional investments, using a sample of 21 ETFs between 2009 and 2013, and shows that while they can effectively beat the markets, their risk-adjusted performance significantly lags behind the conventional ETF benchmarks. In contrast, Chakrabarty et al. (2017) show that SRI ETFs perform at least as well as their benchmark index, and some even outperform the corresponding benchmark. More recently, Pavlova and de Boyrie (2021) study 62 ESG ETFs before and during the Covid-19 crisis induced market crash for six months and document that although ESG ETFs do not protect investors against the crash, they also do not underperform the market.

In light of these mixed findings and the lack of a comprehensive long-term study that explores the potential of value creation from ESG ETFs, it is clear that there is a gap in the literature when it comes to ascertaining long-run benefits from ESG investing using ETFs. Thus, our paper aims to provide further insights into the evolution of ESG ETFs' performance with respect to their benchmarks by studying a sufficiently larger set of ESG ETFs so that their performance can be reliably tracked over the years.

2.3. Passive SRI ETFs: The Hypotheses

The primary purpose of this paper is to examine whether SRI ETFs can generate abnormal returns for investors in comparison to their conventional benchmarks, and to identify the strategies (inclusion or exclusion screening) and attributes (ESG- or product-related) that drive the results. There are many long-run studies examining the effect of SRI on financial performance (e.g., Nofsinger and Varma, 2014), yet none of them focus on ETFs by examining their performance over a prolonged period. Moreover, while existing literature mainly focuses on socially responsible mutual funds, this study examines ETFs and their potential for creating returns for investors by adopting SRI principles. Indeed,

year-on-year growth of both ETFs and SRI give evidence that investors value these assets. On one hand, research has shown that investors pay a price for investing in a socially responsible way (Girard et al., 2007; Renneboog et al., 2008b). On the other hand, there is evidence that integrating SRI into financial investments can help outperform the conventional counterparts (Derwall et al., 2005; Kempf and Osthoff, 2007). Thus, our first hypothesis is aimed at examining the benefits from SRI ETFs by focusing on passively managed ETFs addressing the following question: Can passive ESG ETFs outperform their passive ETF counterparts that do not employ any ESG screening?

Next, we want to examine whether different ESG attributes create value differently and that the different screening methods (exclusion or inclusion) affect these ETFs' financial performance. As discussed in Section 2.1, for SRI funds in general, there is evidence that different socially responsible attributes might have different effects on financial performance (Derwall et al., 2005; Kempf and Osthoff, 2007; Nofsinger and Varma, 2014). Even for the more younger generation of SRI-based financial instruments, we expect similar heterogeneity in performance within ESG ETFs based on the kind of inclusion/exclusion screens they apply and/or other ESG- and product-related attributes. In other words, the value creating potential of passive ESG ETFs may be driven by only some of the SRI-based investment strategies. We essentially hypothesize that there are significant heterogeneities in abnormal returns or alphas based on different ESG attributes and screening strategies used by the ESG ETFs.

Statistically significant abnormal returns for the SRI ETFs may point toward one or both of the following demand and supply effects. On the one hand, there has been an upsurge in demand for sustainability- or SRI-oriented funds over the years. In response to this, rating such as Morningstar Globe, Lipper Fund ESG scores, etc. are becoming increasingly common to help investors make informed investment choice. We know that the availability of such non-return information is important for fund investors' decisions (Dumitrescu and Gil-Bazo, 2016). From SRI funds' perspective, increased scrutiny due to the increased availability of such information, may have incentivized them to become more prudent in their portfolio selection and investment decisions. We conjecture this to

be the main underlying driver of the demand-side performance effect. On the other hand, the increasing number of SRI funds should have had cascading effects on the industry concentration and market competition. While rising competition need not necessarily affect existing funds' management fees (Box et al., 2019), it can still influence market quality and liquidity (Kharma and Eugster, 2021). Thus, recent influx of new SRI ETFs in the market may have benefited their performance, and this is represented by the supply-side effect. For, actively-managed funds, we know that demand-side factors such as investor heterogeneity (Guercio and Reuter, 2014; Dumitrescu and Gil-Bazo, 2018) and preferences (Riedl and Smeets, 2017; Hartzmark and Sussman, 2019) are important. However, since we study passive SRI ETFs, we hypothesize that their under- or out-performance will most likely be explained by supply-side factors than the demand-side ones.

3. Data and Methodology

3.1. SRI ETFs Data

The socially responsible ETFs and their benchmarks' closing prices are from Refinitiv Datastream, formerly known as Thomson Reuters Eikon, while the common risk factors used to obtain abnormal returns are from the Kenneth R. French database. The 11 years sample period begins from January 2010 and ends in December 2020. During this period, we identify a total of 121 U.S. equity ETFs that can be categorised as SRI or ESG ETFs. There is no all-encompassing method or criteria to identify all available SRI-oriented ETFs. Therefore, we rely on multiple searches in the data collection process to obtain a sufficiently expansive universe of funds. Similar to Nofsinger and Varma (2014), we start by searching for funds with an SRI objective from Refinitiv Datastream. More than 6,000 ETFs closing prices are reported in December 2020. From these funds, we then identify the select few ESG ETFs using publicly available lists of SRI ETFs from the "ETF Database." To ensure that no relevant ESG ETFs are missing, we manually check and add additional funds from SocialFunds and the US SIF website. We also account for all of the ETFs in the previous literature (Sabbaghi, 2011; Meziani, 2014; Chakrabarty et al., 2017). The final dataset of the SRI ETF returns include 107,471 daily observation

points over the 11-year sample period. We use this daily time-series data to calculate an equally weighted portfolio return of the 121 ETFs.

In this paper, we focus on equity ETFs that track an index or benchmark of equities. We focus only on U.S. equity ETFs for several reasons. Firstly, the U.S. is the most developed ETF market in the world, and the majority of publicly available ETFs are domiciled there. Second, by focusing on the U.S. stock markets, we avoid any ambiguity in identifying the data for risk factors that are commonly known to affect stock prices (Fama and French, 1993, 2015). And, finally, using a single country setting helps circumvent any cross-country confounding factors. Note that although our sample ETFs could potentially hold global equities, they are listed in U.S. markets. Along with excluding any bond or fixed-income ETFs, commodity ETFs, and currency ETF; we also leave out synthetic ETFs like the inverse and leveraged ETFs. This is due to the fact that their risk profiles are inherently different from those of equity ETFs and may affect their financial performance.

It is important to account for survivorship bias while examining the performance of ETFs or any other type of mutual funds. This bias occurs when ETFs get merged or liquidated due to weak performance or their inability to attract investors. This could potentially result in an overestimation of the historical performance of any given sample since the observed spectrum would not include the funds with weak performance that get weeded out over the years. Essentially, survivorship bias would cause the performance results to be potentially skewed and artificially inflating the performance of the surviving sample (Carpenter and Lynch, 1999). To ensure that our data is free from survivorship bias, we identify 13 SRI ETFs that were liquidated or merged during the sample period and include them in our analysis.

Securities Exchange Commission (SEC) Electronic Data Gathering, Analysis, and Retrieval (EDGAR) system makes all historical and current fund prospectuses publicly available. For this study, it is important to examine the SRI ETFs' fund-specific characteristics and the SRI strategies they employ for creating value for investors. Thus, to do so, we inspect each of the ETFs prospectuses manually and identify what kind of

screening process (Inclusion or Exclusion) or ESG attributes they use in the investment decision-making process and communicate to investors. Lastly, we also obtain important fund characteristics including Total Net Assets (TNA), age (both in days and years), and expense ratios from Refenitiv Datastream, and their Morningstar Globe ratings.

Table I provides a snapshot of all important information for the 121 SRI equity ETFs in our sample. Panel A gives a detailed summary of the sample and how it has evolved over the years. Full sample period starts in January 2010 and ends in December 2020. Along with the number of ETFs, these funds' TNA, age, and expense ratios are also summarized. ESG ETFs are relatively new instruments in the financial markets. The dramatic increase in the number of available SRI-oriented ETF products can be seen in the summary provided in Table I Panel A with the ESG ETF sample growing from 24 in 2010 to 108 in 2020, with 13 funds that either got merged or liquidated. The increase in the number of ESG ETFs is largely concentrated in the second half of the sample (especially after 2016) as is also clearly visible in Figure I (a). This gives us the motivation to examine the recent sub-periods separately. Thus, in our analysis, along with the full sample period, we will focus on two additional sub-periods to capture the effect of ESG ETFs' growth and development: January 2015 to December 2020 and January 2019 to December 2020.

Notably, although the median management fees or expense ratios for SRI ETFs reduces over the years from 0.57% in the beginning of our sample period to 0.40% in the end, it still remains significantly higher than the one for benchmark ETFs that has a median of 0.09%. This could be explained by the fact that while passive non-SRI counterparts may merely follow the S&P500 index, the SRI ETFs' management may need to be compensated for identifying best possible ESG screens to follow. In our analysis, we account for the effect of these costs by evaluating alphas both before and after fees. In Figure I (a), we also show the decreasing trend of median expense ratios over the years along with the increasing trend seen in the number of SRI ETFs. This plot indicates that growing competition among the SRI ETFs, especially after 2015-2016, may be one of the important factors driving the decline in their management fees. The

Table I

Sample Summary

This table summarizes information for the 121 SRI-oriented U.S.-domiciled equity ETFs in our sample. Full Period refers to the entire 11-year period beginning January 2010 and ending December 2020. The time periods Jan-2010, Jan-2015, and Dec-2020 denote the beginning, middle, and end of our sample period. Panel A reports the summary for all the sample ESG ETFs. TNA refers to total net assets under management of all ETFs. Unique fund managers are identified by listing the common fund managers of all the sample ETFs. For the earlier years, the TNA of ETFs and the median total expense ratios were not available in Refinitiv. Panel B summarizes the SRI ETFs on the basis of their screening strategies (Inclusion or Exclusion) and the ESG attributes they use in investment decisions. Any product-related screen refers to product-specific screens like alcohol, tobacco, gambling, weapons, pornography, abortion, etc. Environmental screening strategy focuses on climate change, pollution, environmental sustainability, renewable energy and clean technologies, and clean water. Social screening strategy involves issues such as equality, diversity, racial or gender diversity in company boards, human rights, and community development. The governance screen strategy focuses on corporate governance issues such as independence of directors, executive compensation, and how the company is managed. For the funds employing ESG screen, the focus is on all of the three attributes mentioned above or distinction cannot be made between the three ESG attributes.

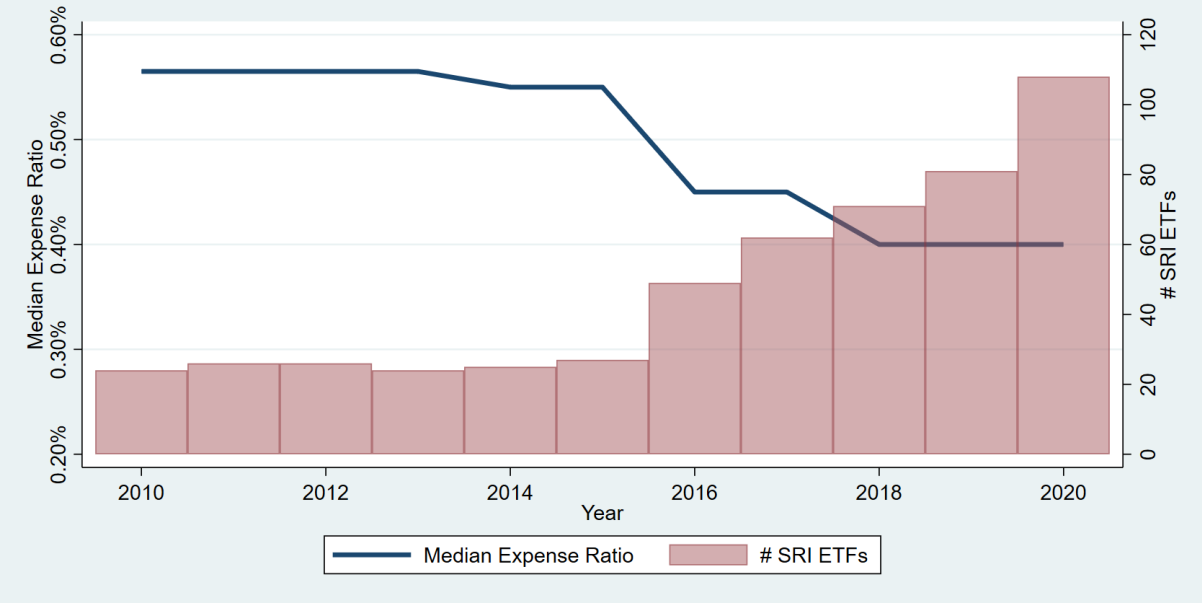
Panel A: All ESG ETFs	Full Period	Jan-2010	Jan-2015	Dec-2020	*% Change (2010 to 2020)*
# of ETFs	121	24	31	108	*350%*
# of unique ETF managers	41	8	13	35	*338%*
Median ETF age		1	2	4	
TNA of all ETFs ($ billions)		3.1	3.8	61.6	*1868%*
Median total expense ratio		0.57%	0.55%	0.40%	*-29%*

Panel B: ESG ETFs by Screening Criteria	Full Period	*% of Total*	Jan-2010	Jan-2015	Dec-2020	*% Change (2010 to 2020)*
Any product-related screen:	54	*45%*	12	15	54	*350%*
ESG screen:						
All (Inclusion or Exclusion)	48	*40%*	3	4	48	*1500%*
Inclusion	25	*21%*	2	2	25	*1150%*
Exclusion	23	*19%*	1	2	23	*2200%*
Environment screens:						
All (Inclusion or Exclusion)	27	*22%*	13	17	27	*108%*
Inclusion	22	*18%*	12	15	22	*83%*
Exclusion	5	*4%*	1	2	5	*400%*
Social screens:						
All (Inclusion or Exclusion)	21	*17%*	2	2	21	*950%*
Inclusion	8	*7%*	1	1	8	*700%*
Exclusion	13	*11%*	1	1	13	*1200%*
Governance screens:						
All (Inclusion or Exclusion)	12	*10%*	2	2	12	*500%*
Inclusion	10	*8%*	1	1	10	*900%*
Exclusion	2	*2%*	1	1	2	*100%*
ESG ETFs (Total):	121	*100%*	24	31	108	*350%*

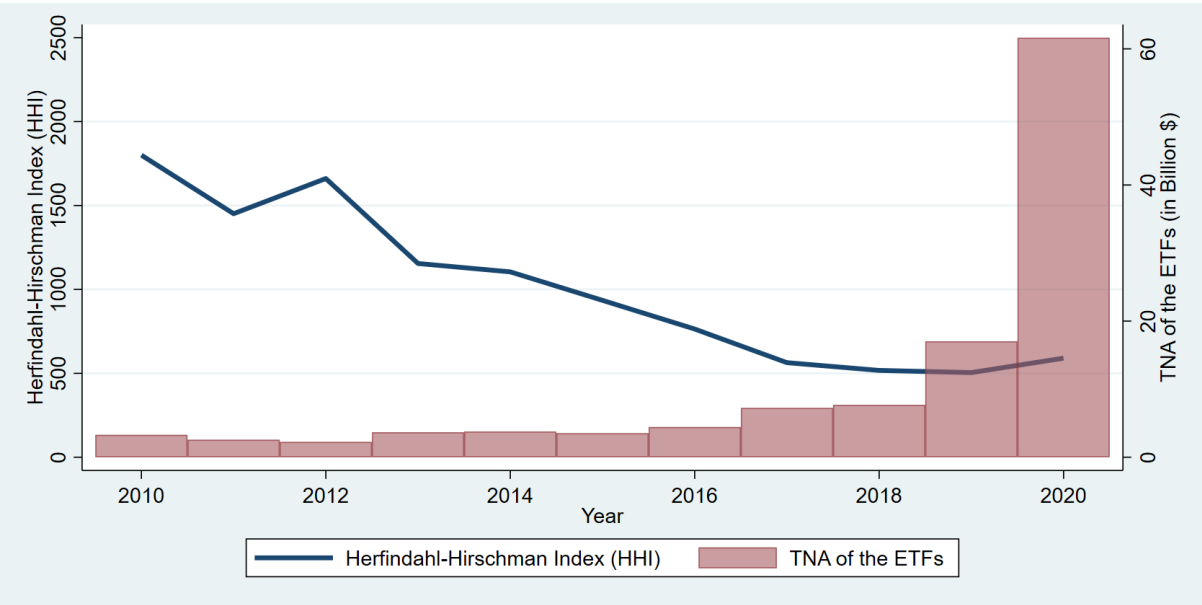
Figure I

Growing Competition Among SRI ETFs Over the Years

This figure provides preliminary evidence of growing competition among the passive SRI equity ETFs during our sample period from 2010 to 2020. The first plot (a) shows the median expense ratios and the number of funds for all the SRI ETFs that we study. The plot of median expense ratio for each year is shown with a line, while the evolution of number of SRI ETFs over the years is represented by bars. Plot (b), meanwhile, shows the trends for the SRI ETFs' total net assets (TNA) and the industry concentrations using Herfindahl–Hirschman Index (HHI). For each year, we identify all the SRI ETFs, their expense ratios, their total TNA, and their industry HHI as on the last trading day of December.



(a) SRI ETFs and their Expense Ratios



(b) TNA of ETFs and the Industry HHI

plot in Figure I (b) compares and contrasts the SRI ETFs' industry concentration (using Herfindahl-Hirschman Index - HHI) over the years with their cumulative totel net assets (TNA). Once again, we see that the HHI falls below 1,000 around 2015-2016 confirming that SRI ETFs market has indeed become less concentrated and more competitive.

Panel B of Table I shows the distribution of all the SRI screening criteria employed by our sample ETFs. The distribution of different SRI categories in these ETFs follow a similar pattern to that shown in Nofsinger and Varma (2014) for SRI mutual funds. SRI ETFs explicitly describe their SRI principles in their fund prospectuses. We find that 45% of the SRI ETFs use a product screen in investment decision making. ESG screen is the second most frequent strategy (40%) and appears to have the fastest growth during our sample period. At the beginning of the sample, environmental screens dominate the industry, with 13 ETFs out of the total 24 employing this strategy. In general, social screens represent the second-largest portion from the single ESG screens, while the governance screen presents the smallest sample in the full sample period. We can also see from the summary table that the ETF managers tend to use more of inclusion screens rather than exclusion in their investment decision process.

3.2. Benchmark (Non-SRI) ETFs Data

To assess whether the SRI ETFs perform better than the conventional non-SRI ETFs, we need to identify a benchmark. Like Cremers et al. (2012) point out, an insufficient choice of benchmarks can cause biased performance results. Practitioners often use a simple benchmark index to measure the performance of mutual funds and ETFs. A suitable benchmark, thus, should be a passively managed portfolio of ETFs with similar risk exposures to evaluate the added value from our passive ESG ETF portfolio.[5]

In this study, we compare SRI ETFs to four different S&P500 equity ETFs representing the benchmark market index. We include four different S&P500 equity ETFs from different issuers to mitigate any effect of tracking errors or incorrectly timed equity purchases. The final dataset of the benchmark ETFs has 11,080 daily observation points

[5]Some studies use the standard asset pricing models like the Fama and French (1993) three-factor model and Carhart (1997) four-factor model as benchmarks, but for passive indexes, these may result in significant negative alphas.

over the 11-year sample period. We use these to compute daily time-series of returns with an equally weighted portfolio of the four ETFs. Following Nofsinger and Varma (2014), we create an additional difference portfolio between the SRI ETFs and the benchmark portfolios. Similarly, for each of the SRI screening strategy-based portfolios, difference portfolio is created by subtracting the returns of the benchmark ETFs portfolio from the different screening-based portfolios.

3.3. Portfolio Returns

Table II shows the detailed summary statistics for the daily returns of each equally weighted portfolio that we analyse.[6] SRI ETFs portfolio is the total sample of 121 SRI ETFs, S&P500 ETFs portfolio is an equal weighted portfolio of four different S&P500 ETFs, and the difference portfolio is the difference between these two portfolios. Finally, the remaining portfolios reported in the table are used to examine the performance of different screening strategies and ESG attributes. All of the portfolios are yielding positive returns on average during the sample period. Mean and median returns are relatively close to each other. Note that these statistics are using daily returns and the effect widens with monthly or yearly observations. Of all the screening strategies, Environmental Inclusion strategy provides the highest daily return while the S&P500 ETF portfolio has the lowest daily return during the period. Environmental Inclusion also exhibits the highest volatility during the sample period.

3.4. Empirical Approach

3.4.1. Estimating the Abnormal Returns

To test the performance of SRI ETFs over the passive non-SRI counterpart, we use Carhart (1997) four factor model along with the CAPM model, the Fama and French (1993) three-factor model, and the Fama and French (2015) five-factor model for robustness checks. It is common in the literature to use similar asset pricing models for SRI-based funds (Bauer et al., 2005; Derwall et al., 2011). Since SRI ETFs are a recent innovation within SRI-based mutual funds, for ease of comparison with the past literature

[6] All the portfolio characteristics remain qualitatively similar with value weighted returns and so do our main results. So, they are left out for brevity.

Table II

Portfolio Returns

This table presents the descriptive statistics of the daily returns for each portfolio expressed in percentage units. Along with the mean, median, and standard deviation (SD), we report the extreme values (i.e., Min and Max) and higher order moments (kurtosis and skewness represented by Kurt and Skew respectively). SRI ETF portfolio comprises of the full sample of 121 SRI ETFs, S&P500 ETF portfolio is formed using four different S&P500 ETFs, and the difference portfolio is the difference between the two portfolios. The rest of the portfolios are shown to summarize the performance of different screening strategies and ESG attributes. All of the portfolios are equally weighted.

Portfolios	# ETFs	Min	Max	Mean	Median	SD	Kurt	Skew
SRI ETF portfolio	121	-0.1064	0.0851	0.0003	0.0008	0.0120	8.4617	-0.6362
S&P500 ETF portfolio	4	-0.1148	0.0939	0.0005	0.0007	0.0108	15.5856	-0.6287
Difference portfolio	-	-0.1064	0.0851	0.0003	0.0008	0.0121	8.3149	-0.6189
Inclusion	67	-0.1088	0.0903	0.0003	0.0009	0.0124	8.4624	-0.6184
Exclusion	38	-0.1065	0.0837	0.0005	0.0009	0.0104	13.0473	-0.8018
Product Screen	54	-0.1072	0.0878	0.0005	0.0008	0.0104	13.8627	-0.7539
ESG Inclusion	25	-0.1084	0.0861	0.0004	0.0008	0.0099	16.3035	-0.9186
ESG Exclusion	23	-0.1041	0.0880	0.0005	0.0008	0.0103	13.5469	-0.7367
Environmental Inclusion	22	-0.1103	0.0975	0.0004	0.0008	0.0130	8.8165	-0.6261
Environmental Exclusion	5	-0.1094	0.0803	0.0004	0.0008	0.0102	12.9719	-0.7553
Social Inclusion	8	-0.1074	0.0897	0.0005	0.0008	0.0106	12.6835	-0.6569
Social Exclusion	13	-0.1070	0.0823	0.0004	0.0008	0.0105	12.7855	-0.7738
Governance Inclusion	10	-0.1036	0.0860	0.0005	0.0008	0.0104	12.9246	-0.6780
Governance Exclusion	2	-0.1078	0.0911	0.0005	0.0008	0.0104	13.3187	-0.7132

that has studied SRI mutual funds, we employ Carhart (1997) four-factor model as the main empirical model:

$$R_{i,t} = \alpha + \beta_1 * RMRF_t + \beta_2 * SMB_t + \beta_3 * HML_t + \beta_4 * MOM_t + \epsilon_{i,t}, \tag{1}$$

where $R_{i,t}$ is the excess return for portfolio i over the risk-free rate for each day t, the three Fama-French factors are represented by $RMRF_t$ (market factor), SMB_t (size factor), and HML_t (book-to-market or value factor), and additional momentum factor is captured by MOM_t. We test the performance of SRI ETFs over three different time periods to understand how the recent upsurge in SRI ETFs has affected their performance: a) full period (January 2010 to December 2020), b) half period (January 2015 to December 2020), and c) last 2 years (January 2019 to December 2020). Thus, the recent time periods are aimed at isolating the impact of growing and maturing SRI ETF market. As previously summarized in Tables I and II, SRI ETFs can implement a variety of screening strategies. To assess the performance of these screening criteria, we use their respective excess returns in place of $R_{i,t}$.

3.4.2. Explaining the Abnormal Returns

To understand how the SRI ETFs' demand and supply dynamics affect their performance, we run empirical tests at both portfolio and fund levels. In the fund-level analysis, we track all the SRI ETFs' Morningstar Globe ratings from March 2016 when the first set of ratings were made available and their relative industry concentration or market power. To do so, we estimate the following model using each of the SRI ETFs' excess returns:

$$\begin{aligned} Excess\ Returns_{j,t} = \gamma_0 + \gamma_1 Globe\ Rating_{j,t} + \gamma_2 Concentration_{j,t} \\ + \gamma_3 Fund\ Characteristics_{j,t} + \gamma_5 Screening_{j,t} + u_{j,t}, \end{aligned} \tag{2}$$

where the $Excess\ Returns_{j,t}$ is the daily excess returns of fund j for day t, and the main independent variables are $Globe\ Rating_{j,t}$, which is the ETF j's available Morningstar Globe rating for that day t, and $Concentration_{j,t}$, which is ETF j's relative market concentration. Following Renneboog et al. (2008a) and Hartzmark and Sussman (2019), we control for $Fund\ Characteristics_{j,t}$ that includes size (log of TNA), expense ratio,

age (log of days since inception), and last 20 days returns and return volatility (roughly accounting for one month). $Screening_{j,t}$ is an array of SRI screens used by the ETFs (product, inclusion, exclusion, etc.). In the baseline model, we also control for day fixed effects to account for time-specific trends and any macroeconomic events or market-wide shocks. This specification captures both the between- and within-ETFs impact of our two main explanatory variables on the variation in their returns. In addition, we estimate two alternative specifications of the baseline model by a) including fund fixed effects (to isolate the within-ETFs impact) and b) using Fama-MacBeth two-step approach (to assess the between-ETFs effect).

Since all our main sample ETFs belong to the same industry, using an industry-wide measure of market concentration such as HHI is problematic as it does not capture relative market power of each fund. Thus, to measure fund-specific market concentration, first we obtain each SRT ETF's market share for each day using its TNA (i.e., $Share_{j,t} = TNA_{j,t} / \sum_{i=1}^{n} TNA_{i,t}$), and then compute the average market share for SRI ETFs excluding the focal fund:

$$Concentration_{j,t} = \frac{\sum_{\substack{i=0 \\ i \neq j}}^{n} Share_{i,t}}{n-1}. \tag{3}$$

In essence, this measure of relative market power or competition is very similar to the measure of peer pressure used in Cao et al. (2018).

In the portfolio-level analysis, we follow Hartzmark and Sussman (2019) and form portfolios of our sample SRI ETFs each day by classifying them based on their Morningstar Globe ratings and them obtain their respective equal-weighted daily returns. Thus, 1-Globe rated portfolio tracks the returns for low SRI- or sustainability-orientation, while the 5-Globes portfolio represents the high sustainability ETFs. Along with the five Globe ratings-based portfolios, we also measure the difference between two extreme portfolios to trace a long 5-globes - short 1-globe hedging strategy. Finally, for each of these portfolios, we estimate abnormal returns or alphas using the four-factor model in Equation (1) and other alternative asset pricing models.

4. Performance of SRI vs Benchmark ETFs

Our baseline results estimate the performance of passive SRI ETFs and their benchmark non-SRI ETFs separately along with their difference portfolio for different time periods. Following Bauer et al. (2005), we estimate the alphas for each of these portfolios before and after deducting management fees. To obtain the pre-fees alphas, we add back management fees to the ETF returns before estimating the Carhart (1997) four-factor model.

The results for both post- and pre-fees alphas are shown in Table III Panels A and B respectively using the four-factor model. Over the full sample period, the SRI ETF portfolio performs worse than the benchmark portfolio of passive S&P500 ETFs. The SRI ETF portfolio (non-SRI benchmark portfolio) has an annualized alpha of -7.32% (-1.87%). Meanwhile, the difference portfolio that captures the SRI-based premium provides a negative annualized alpha of -5.35%. All these alphas are statistically significant at 1%. During the latter half period, the SRI ETFs begin to outperform the benchmark portfolio of passive S&P500 ETFs. While the SRI ETFs generate 0.08% annualized alpha, the benchmark non-SRI portfolio generates statistically significant alpha of -2.34%. Interestingly, for the final two years of our sample, the outperformance of the SRI ETFs becomes even more visible with an annualized alpha of 5.25% ($p \leq 0.10$) when compared to the benchmark ETFs that show an alpha of -2.64% ($p \leq 0.05$).

Even when we account for the management fees charged by both the SRI and non-SRI ETFs in Table III Panel B, the SRI ETF portfolio underperforms the benchmark portfolio of passive S&P500 ETFs over the full sample period. The SRI ETF (non-SRI benchmark) portfolio has a statistically significant annualized alpha of -6.75% (-1.79%) when the ETF returns are adjusted for the management fees. The difference portfolio has a slightly lower alpha of -4.96% when compared to that using post-fees returns due to the higher fees charged by the SRI ETF funds. Similar to Panel A, in the second half of the sample period, and especially in the last two years, the SRI ETF portfolio outperforms the benchmark portfolio. In fact, between January 2019 and December 2020, the difference portfolio annualized alpha increases to 8.28%, which is statistically significant at 5%.

Table III

Performance of SRI vs Benchmark ETFs

This table reports the four-factor model alphas for the sustainable SRI ETF portfolio, the benchmark S&P500 ETF portfolio, and the difference portfolio that tracks a long SRI ETFs short non-SRI S&P500 ETFs hedge. Alphas are estimated using equally-weighted portfolios of the ETFs from 2010 to 2020. Three different time periods are shown: Full Period (January 2010 to December 2020), Half Period (January 2015 to December 2020), and Last 2 Years (January 2019 to December 2020). All alphas are annualized for ease of interpretation and expressed in percentage. Panel A reports the post-fees portfolio alphas, while Panel B reports the pre-fees alphas. The pre-fees alphas are calculated by first adjusting each of the ETF returns for their management fees, and then estimating the four factor model. Significance levels are shown as follows: *** at 1%, ** at 5%, and * at 10%. The reported t-stats are given in the brackets.

Panel A: Post-fees Alphas Time Periods:	Full Period	Half Period	Last 2 Years
# of SRI ETFs	*121*	*115*	*114*
SRI ETF Portfolio	-7.32***	0.08	5.25*
	[-3.37]	[0.03]	[1.70]
S&P500 ETF Portfolio	-1.87***	-2.34***	-2.64**
	[-2.90]	[-3.55]	[-2.48]
Difference Portfolio	-5.35***	2.33	8.02**
	[-2.46]	[1.02]	[2.24]
Panel B: Pre-Fees Alphas Time Periods:	Full Period	Half Period	Last 2 Years
# of SRI ETFs	*121*	*115*	*114*
SRI ETF Portfolio	-6.75***	-0.59	5.75*
	[-3.34]	[-0.28]	[1.85]
S&P500 ETF Portfolio	-1.79***	-2.24***	-2.53**
	[-2.78]	[-3.43]	[-2.35]
Difference Portfolio	-4.96***	1.66	8.28**
	[-2.41]	[0.77]	[2.52]

During this period, the SRI ETF portfolio (benchmark non-SRI portfolio) yields an alpha of 5.75% (-2.53%) that is statistically significant at 10% (5%).

For the full sample period, Table IV shows the factor loadings and adjusted R^2 values for the estimated Carhart (1997) four factor model. Since we focus on passive ETFs that track the market portfolios, it is not surprising at all that the estimated models show very high explanatory power. The benchmark non-SRI portfolio expectedly has the highest adjusted R^2 of 0.992, while the SRI ETF portfolio has an adjusted R^2 of 0.932. The market beta coefficients (for $RMRF$) are relatively close to each other (0.95 and 0.96), signifying that the SRI ETFs and the benchmark returns move in the same direction as the market but are bit more defensive and less diversified. The size factor (SMB) coefficient is positive for the SRI ETF portfolio and the difference portfolio, but is negative for the benchmark ETFs portfolio. This shows that there is a distinctively opposite correlations for SRI and benchmark ETFs with the size factor. When compared to the benchmark S&P500 ETFs portfolio, SRI ETFs' portfolio returns load significantly more on the book-to-market (HML) and momentum (MOM) factors. Notably, we find that all of the coefficients of the difference portfolio (i.e., long SRI ETFs short benchmark ETFs hedge) closely track the coefficients of the SRI ETF portfolio. This can be explained by the fact that much of the variations in benchmark portfolio returns are largely explained by the four factor model (as is also visible by its high R^2 value).

Figure II shows the cumulative abnormal returns using the estimated mean daily abnormal returns from the four factor model estimated over 3-months rolling window. For each day, the average abnormal returns are estimated over the next 60 trading days, and then their cumulative sums are computed over the three periods summarized in Tables III. All the three plots in the figure are consistent with the previous results. Over the full sample period (a), the cumulative abnormal returns for the SRI ETF portfolio decrease and stay below that of the benchmark S&P500 portfolio through the years. However, when plotted for the half period (b), the SRI and benchmark portfolios' returns slowly start diverging with those of the SRI ETFs remaining flat and hovering around zero while the S&P500 ETFs' remains negative. Due to this, the difference portfolio shows a

Table IV

Factor Loadings for SRI vs. Benchmark ETFs

This table reports the four-factor model coefficient estimates for the sustainable SRI ETF portfolio, the benchmark S&P500 ETF portfolio, and the difference portfolio that tracks a long SRI ETFs short non-SRI S&P500 ETFs hedge. We report the summary of estimations for the full sample period (January 2010 to December 2020) using both post-fees portfolio returns (Panel A) and pre-fees portfolio returns (Panel B). The pre-fees ETF returns are obtained by adjusting each of the reported (or, post-fees) ETF returns for their management fees. Similar to Table III, all alphas are annualized for ease of interpretation and expressed in percentage. Significance levels are shown using ***, **, and * at 1%, 5%, and 10% respectively. The t-stats for each of the coefficients are reported in the brackets.

Panel A: With post-fees returns	*Alpha*	*RMRF*	*SMB*	*HML*	*MOM*	*Adj.*R^2
SRI ETF Portfolio	-7.32***	0.95***	0.20***	0.07***	-0.04**	0.932
	[-3.37]	[122.46]	[12.98]	[4.53]	[-3.31]	
S&P500 ETF Portfolio	-1.87***	0.96***	-0.13***	0.00	-0.01**	0.992
	[-2.90]	[402.79]	[-27.84]	[0.95]	[-2.57]	
Difference Portfolio	-5.35***	0.95***	0.20***	0.06***	-0.03***	0.927
	[-2.46]	[118.00]	[12.39]	[4.01]	[-2.64]	
Panel B: With pre-fees returns	*Alpha*	*RMRF*	*SMB*	*HML*	*MOM*	*Adj.*R^2
SRI ETF Portfolio	-6.75***	0.92***	0.19***	0.08***	-0.03***	0.932
	-[3.34]	[122.23]	[12.53]	[5.33]	-[2.91]	
S&P500 ETF Portfolio	-1.79***	0.96***	-0.13***	0.005	-0.0090**	0.992
	-[2.78]	[402.79]	-[27.84]	[0.95]	-[2.57]	
Difference Portfolio	-4.96***	0.92***	0.18***	0.07***	-0.0255**	0.926
	-[2.41]	[117.49]	[11.93]	[4.75]	-[2.23]	

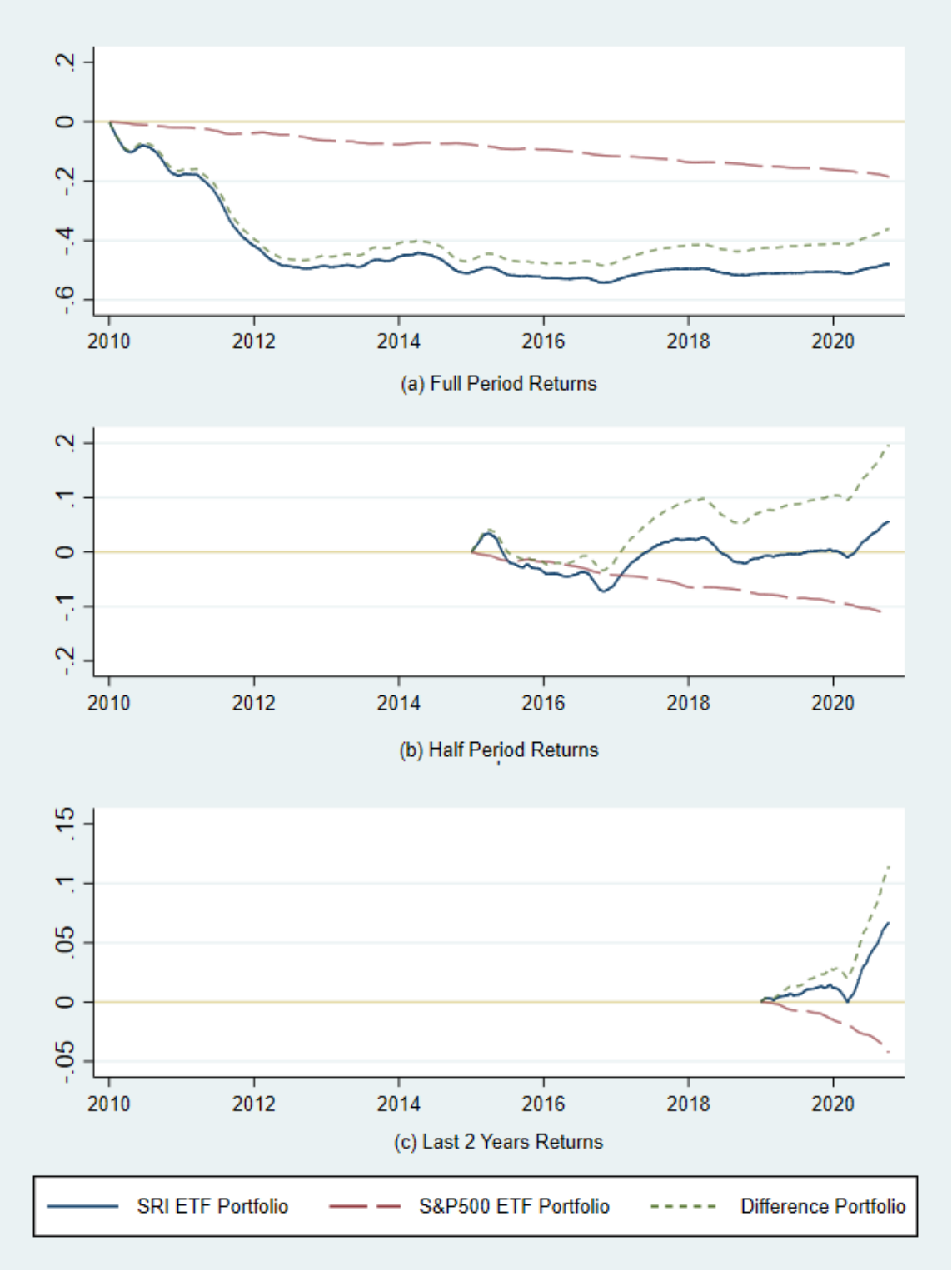
potential to generate positive abnormal returns. Meanwhile, for the last two years (c), the benefits from SRI ETFs are more clearly visible as the underperformance of benchmark portfolio is evident.

We alternatively re-estimate all our results that employ the Carhart (1997) four-factor model using three other asset pricing models and find that they remain consistently robust throughout. In place of the four-factor model, we apply a) the five-factor model (Fama and French, 2015), b) the three-factor model (Fama and French, 1993), and c) the capital asset pricing model (CAPM). These results are shown in the Appendix A Tables A.I and A.II. Among these three models, the alpha for CAPM is the highest in terms of magnitude for both the SRI ETF portfolio and the difference portfolio, whereas the Fama-French five factor model gives the most conservative estimates of alphas.

Figure II

Cumulative Abnormal Returns from SRI vs. Benchmark ETFs

This figure shows the daily cumulative excess returns for the SRI ETF portfolio, the benchmark S&P500 ETF portfolio, and the Difference portfolio that mimics a long short hedge of the first two portfolios that is rebalanced daily. We compute the cumulative excess returns as follows. For each day, the average daily abnormal return is computed using a 3-month (or, 60 trading days) rolling window. Similar to Tables III and IV, in these estimations, we employ the four factor model. Then, the cumulative sum of the estimated abnormal returns are computed on a daily basis for (a) Full Period starting the first trading of January 2010, (b) Half Period beginning in January 2015, and (c) Last 2 Years of the sample period starting January 2019.



5. Performance of Screening Strategies

So far, the performance analysis has analyzed SRI ETFs as a homogeneous group. However, different screening strategies of the SRI ETFs may perform differently. Thus, next, we separate the SRI ETFs sample into sub-samples based on the strategies that they employ in their investments. The objective here is to examine which of the strategies (inclusion or exclusion) and attributes (ESG and product-related) drive the results observed for the SRI ETF portfolio in the previous section. Any product-related screen refers to an exclusion screen where the ETF avoids certain "sin" industries' products like alcohol, tobacco, gambling, weapons, etc. An inclusion strategy involves positive screening that employs best-in-class approach where the ETF overweights assets that perform well with on either some specific or all ESG attributes. Alternatively, under the exclusion strategy ETF leaves out assets that perform poorly on one or all ESG attributes.

Table V summarizes the performance of all the SRI screening strategies used by the SRI ETFs over three different periods similar to Table III, i.e., full sample period, second half of the sample, and last two years. Over the full sample period, while the Exclusion strategy neither under- or over-performs the market, the inclusion strategy yields a statistically significant annualized alpha of -6.30% ($p \leq 0.01$). Much of the underperformance of the inclusion strategy appears to be driven by the ETFs that incorporate the Environmental Inclusion strategy, with the portfolio of these ETFs generating an annualized alpha of -5.61% ($p \leq 0.05$). Note that the lack of statistical significance for some of the estimated alphas may simply be due to the small subsample size for some SRI-based screens. For example, there are only two ETFs in our sample that use Governance Exclusion in their investment selection. These trends are similar to those observed in Nofsinger and Varma (2014).

As shown in Table III, the returns from SRI ETFs start to turn positive in the latter part of the sample period. It is also the time period during which many of the ETFs were launched, and both SRI and ETFs independently gained significant market traction. In the recent half period of our sample (2015–2020), both the Inclusion- and Exclusion-based ETFs have statistically insignificant positive annualized abnormal returns, meaning that

Table V

Performance of SRI ETFs' Screens

This table reports the four-factor model alphas by dividing the sample SRI ETFs on the basis of their screening strategies (Inclusion or Exclusion) and the ESG attributes they use in investment decisions. Product screen refers to any product-specific screening that avoids industries like alcohol, tobacco, gambling, weapons, pornography, abortion, etc. The environmental inclusion/exclusion strategy focuses on climate change, pollution, environmental sustainability, renewable energy and clean technologies, and clean water. Social inclusion/exclusion strategy focuses on issues such as equality, diversity, racial or gender diversity in company boards, human rights, and community development. The governance inclusion/exclusion strategies focus on issues like independence of directors, executive compensation, and how the company is managed. Under the ESG inclusion/exclusion screening, either a combinations of all of the three ESG attributes are integrated or cannot be specifically determined. Similar to Tables III and IV, all alphas are annualized for presentation purposes and expressed as percentage. Significance levels are represented by: *** at 1%, ** at 5%, and * at 10%. The reported t-stats are provided in the brackets. All the strategies are implemented using the daily time-series of an equally-weighted ETF portfolio returns for the full sample period 2010–2020.

Time Period:	Full Period	Half Period	Last 2 Years
SRI Screening Portfolios:			
Inclusion	-6.31***	0.35	7.81**
	[-3.01]	[0.15]	[2.42]
Exclusion	-0.30	0.66	0.43
	[-0.20]	[0.32]	[0.15]
Product Screen	-1.56	-1.33	-2.12
	[-1.28]	[-0.94]	[-0.80]
ESG Inclusion	-1.57	-2.45	-2.12
	[-1.09]	[-1.28]	[-0.80]
ESG Exclusion	-1.10	-0.66	0.74
	[-0.73]	[-0.31]	[0.30]
Environmental Inclusion	-5.61**	3.46	13.80***
	[-2.20]	[1.10]	[3.63]
Environmental Exclusion	-2.23	-2.63	-1.88
	[-1.54]	[-1.31]	[-0.60]
Social Inclusion	-1.77	-1.55	-0.80
	[-1.36]	[-0.95]	[-0.33]
Social Exclusion	-2.04	-2.28	-0.22
	[-1.38]	[-1.11]	[-0.07]
Governance Inclusion	-1.37	-0.97	-0.27
	[-1.02]	[-0.56]	[-0.12]
Governance Exclusion	-1.90	-1.84	-1.25
	[-1.48]	[-1.16]	[-0.40]

neither of the strategies perform worse than the markets. During this period, none of the ESG based inclusion or exclusion strategies generate statistically significant alphas.

In the last two years of our sample period, there is significant outperformance by the Environmental Inclusion screening strategy as it yields 13.80% annualized alpha ($p \leq 0.01$), clearly driving the results for both the Inclusion portfolio (alpha of 7.81% at $p \leq 0.05$) and the overall SRI ETF portfolio (see Table III). This ability of environmental ETF screens to generate abnormal returns is consistent with the findings in the past literature for individual stock-based portfolios (Kempf and Osthoff, 2007) as well as for SRI-based mutual fund portfolios in general (Nofsinger and Varma, 2014). Across all the three time periods, the product screening strategy marginally underperforms the markets, but for this portfolio alphas are statistically insignificant throughout. These results are indicative of exclusion strategy being associated with increased investment risk that can negatively affect returns (Humphrey and Tan, 2014).

Table VI, meanwhile, presents the factor loadings for all of the different SRI screening portfolios. The product screening strategy by ETFs results in returns that report the highest adjusted R^2 of 0.969 for the four-factor model, while the lowest adjusted R^2 is observed for environmental inclusion. With the market factor loading close to 1, both broad-based inclusion and environmental inclusion strategies show the highest ability of tracking the market movements.

In general, our results from the performance evaluation of the different SRI strategies clearly demonstrate their heterogeneous affect on portfolio returns. We find that it is mainly the ETFs that use Environmental Inclusion as a screening strategy that show the potential to generate significant abnormal returns. Moreover, there is a marked difference in the ability of this portfolio to underperform the markets over the full sample period in comparison to the last two years of our sample when it outperforms the markets.

6. What Explains the SRI ETFs performance?

So far our analysis focused on measuring and tracking passive SRI ETFs' performance in comparison to their benchmarks, while also assessing different SRI screening strategies that they employ. In this section, we examine the factors influencing SRI

Table VI

Factor Loadings for SRI ETFs' Screens

This table reports the four-factor model coefficient estimates when the sample SRI ETFs are divided on the basis of their screening strategies (Inclusion or Exclusion) and the ESG attributes they use in investment decisions. Product screens do not consider sin industries' stocks for investments. The environment, social, and governance inclusion/exclusion strategy focuses on each of the ESG-based characteristics to select investments. We report the summary of estimations for the full sample period (January 2010 to December 2020). Similar to Table III, all alphas are annualized for ease of interpretation and expressed in percentage. Significance levels are depicted in the table using ***, **, and * for 1%, 5%, and 10% respectively. The t-stats for each coefficient is reported in the brackets underneath.

	Alpha	*RMRF*	*SMB*	*HML*	*MOM*	*Adj.*R^2
SRI Screening Porfolios:						
Inclusion	-6.31***	0.99***	0.23***	0.07***	-0.03**	0.936
	[-3.01]	[126.78]	[14.70]	[4.51]	[-2.37]	
Exclusion	-0.30	0.88***	0.03***	0.03***	-0.02***	0.956
	[-0.20]	[160.13]	[3.06]	[3.04]	[-2.94]	
Product Screen	-1.56	0.89***	0.00	0.02**	-0.02***	0.969
	[-1.28]	[195.73]	[-0.27]	[2.23]	[-3.51]	
ESG Inclusion	-1.57	0.84***	-0.01***	0.04***	-0.03***	0.954
	[-1.09]	[156.64]	[-0.85]	[3.85]	[-3.54]	
ESG Exclusion	-1.10	0.87***	0.03**	0.03***	-0.03***	0.953
	[-0.73]	[154.35]	[2.55]	[2.71]	[-3.17]	
Environmental Inclusion	-5.61**	0.99***	0.33***	0.12***	-0.01	0.913
	[-2.20]	[104.30]	[17.48]	[6.26]	[-0.43]	
Environmental Exclusion	-2.23	0.88***	-0.03***	-0.01	-0.01	0.956
	[-1.54]	[162.65]	[-3.12]	[-0.96]	[-1.46]	
Social Inclusion	-1.77	0.91***	-0.03***	0.04***	-0.03***	0.967
	[-1.36]	[187.57]	[-3.10]	[3.77]	[-4.49]	
Social Exclusion	-2.04	0.88***	0.04***	0.04***	-0.03***	0.957
	[-1.38]	[161.40]	[3.47]	[3.45]	[-3.80]	
Governance Inclusion	-1.37	0.89***	-0.04***	0.02**	-0.03***	0.963
	[-1.02]	[177.85]	[-3.75]	[2.40]	[-3.61]	
Governance Exclusion	-1.90	0.90***	-0.03***	0.00	-0.01*	0.967
	[-1.48]	[188.99]	[-3.38]	[-0.04]	[-1.86]	

Table VII

Performance Determinants of SRI ETFs

This table shows the estimation of Equation (2) to study the determinants of SRI ETF returns. In columns (2), (4), and (6), ETFs' fund characteristics and screening strategies are included as control variables (for definitions, see Section 3.4.2). However, many of these time-invariant fund characteristics are omitted from the fund fixed effects estimations in (4). Last two columns show the results from Fama-Macbeth two-stage estimation. The funds' Globe Ratings range from 1 to 5 with 5-Globes rating indicative of high sustainability ETFs. In the results reported here, all ETFs are assigned 0 Globe for days before the availability of these ratings by Morningstar (i.e., before March 2016). Significance levels are shown using *** at 1%, ** at 5%, and * at 10% for each of the reported coefficients and their respective robust t-stats are provided in the brackets.

	Baseline		Fund Fixed Effects		Fama-MacBeth	
	(1)	(2)	(3)	(4)	(5)	(6)
Globe Rating	-0.0035	-0.0046	-0.0108	-0.0263***	-0.0011	-0.0001
	[-1.09]	[-1.33]	[-1.35]	[-3.02]	[-0.96]	[-0.04]
Concentration	-3.3348***	-1.8956*	-4.7741***	-4.6422***	-9.4380***	-3.4170*
	[-3.39]	[-1.84]	[-2.84]	[-2.66]	[-3.03]	[-1.96]
Fund Characteristics	No	Yes	No	Yes	No	Yes
Screening	No	Yes	No	Yes	No	Yes
Time Fixed Effect	Yes	Yes	Yes	Yes	No	No
Fund Fixed Effect	No	No	Yes	Yes	No	No
N	107,471	92,664	107,471	92,664	107,471	92,664
Adj. R^2 / **Avg.** R^2	0.698	0.730	0.699	0.731	**0.0455**	**0.391**

ETFs performance. As investors' demand for sustainability-oriented funds (and related fund flows) is known to be significantly influenced by independent third-party ratings (Hartzmark and Sussman, 2019), we capture the demand-side effect using the ETFs' Morningstar Globe ratings. In contrast, the supply-side effect is reflected in the relative market concentration faced by each of the ETFs over time.

Table VII reports the results obtained from estimating Equation (2) to explore and understand two possible drivers of SRI ETF performance: their sustainability ratings (*Globe Ratings*) and the market competitiveness that they face (*Concentration*). First two columns report the coefficients when day fixed effects are controlled for, next two columns show the results when both day and fund fixed effects are included, and the last two columns show the coefficients from Fama and MacBeth (1973) two-stage cross-sectional regressions. Across all of the different estimation models, we see consistently strong evidence that increase in market concentration (proxied by the average market share of competitors) has a negative impact on ETF returns. For ease of interpretation, all the coefficients representing the sensitivity of the variables to excess returns are

Table VIII

Performance of Globe Ratings-sorted SRI ETFs

This table reports the alphas using multiple asset pricing models when the sample SRI ETFs are classified on the basis of their Morningstar Globe Ratings. The sample starts in March 2016 when the Globe Ratings were first released. A fund with 1-Globe rating represents low sustainability-orientation, while a fund with 5-Globes rating is classified as high sustainability ETF. Along with the portfolio returns of five different Globe ratings-based SRI ETF portfolios, we also measure a difference portfolio for the two extremes (i.e., long 5-globes / short 1-globe zero investment hedge). All alphas are annualized for presentation purposes and expressed in percentages. Significance levels are represented by: *** at 1%, ** at 5%, and * at 10%. The reported t-stats are provided in the brackets. All the strategies are implemented using a daily time-series of equally weighted returns with the sample period ending in December 2020.

	1 Globe	2 Globes	3 Globes	4 Globes	5 Globes	5–1 Hedge
Carhart Four Factors Model	-1.6402***	-1.8253***	-1.6735***	-1.7971***	-1.7073***	-0.0255
	[-8.90]	[-12.07]	[-10.89]	[-12.29]	[-10.74]	[-0.45]
Fama-French 5 Factors Model	-1.6007***	-1.7971***	-1.6402***	-1.7691***	-1.6802***	-0.0307
	[-8.72]	[-11.91]	[-10.72]	[-12.13]	[-10.62]	[-0.55]
Fama-French 3 Factors Model	-1.6072***	-1.7971***	-1.6468***	-1.7691***	-1.687***	-0.0307
	[-8.75]	[-11.94]	[-10.75]	[-12.15]	[-10.64]	[-0.53]
Capital Asset Pricing Model	-1.6336***	-1.8182***	-1.6735***	-1.8041***	-1.7141***	-0.0307
	[-8.88]	[-12.10]	[-10.88]	[-12.29]	[-10.79]	[-0.55]

expressed in basis points. In contrast, the availability and magnitude of sustainability ratings does not have statistically significantly association with SRI ETF returns (except in column (4)).[7] This finding is consisted with Hartzmark and Sussman (2019) that shows sustainability ratings for mutual funds only affects their fund flows but not their performance. Overall, we observe stronger evidence suggesting that supply-side factors like competition concentration are more influential to SRI ETFs returns than the demand-side factors.

We further run complementary tests using portfolio-level analysis. In this test, we divide all the SRI ETFs in Morningstar Globe-based portfolios and then estimate their risk-adjusted returns or alphas using Equation I. The results are shown in Table VIII, which again confirm the findings from Table VII that SRI ETFs' sustainability ratings are not related to their returns.

[7] Our results remain robust to an alternative estimation where the sample size is restricted only to the period when the Morningstar Globe ratings are available during our sample years i.e., March 2016 to December 2020.

7. Conclusion

We examine the potential benefits for investors from a relatively new investment vehicle that adopts SRI principles in ETFs. With both the ETFs and SRI funds independently gaining traction among investors in recent years, ESG ETFs have seized this opportunity to significantly grow their assets under management. Thus, it is important to understand whether these SRI ETFs simply serve investors' altruistic needs at additional financial costs or can also simultaneously generate abnormal returns for the investors. To do so, we exploit a unique survivorship-free dataset of 121 passive U.S. equity SRI ETFs from January 2010 to December 2020, and compare the performance of SRI ETF portfolio with a benchmark portfolio consisting of passive S&P500 ETFs.

Over the full sample period of 11 years, the passive S&P500 ETFs performs better than the passive SRI ETFs. The difference portfolio of SRI ETFs and its passive benchmark has a statistically significant negative abnormal returns of 5.35%, despite both the benchmark S&P500 ETFs and the SRI ETFs underperforming the market. However, the benefits from SRI ETFs begin to appear in the latter half of the sample period with the SRI ETFs performing better than the S&P500 ETFs. Between 2015 and 2020, the SRI ETF portfolio yields 0.08% annualized alpha while the benchmark ETF portfolio has an alpha of -2.34%. Furthermore, when we limit the investment period to the last two years (January 2019 to December 2020), the SRI ETF portfolio outperforms the market and generates 5.25% annualized abnormal returns. In contrast, the benchmark portfolio underperforms the market yielding -2.64% annual alpha. This provides evidence that passively managed SRI ETFs are not necessarily consistently losing to the passive ETF benchmarks. Pavlova and de Boyrie (2021) study SRI ETFs for six months around the onset of Covid-19 crisis and find that while they do not provide a safety net against market crashes, they also do not perform worse than the markets. In this regard, our results showing SRI ETF portfolio outperformance in the last two years indicate that despite their inability to beat the markets during crisis period, the SRI ETF investors could potentially still have reaped benefits in comparison to their non-SRI counterparts.

Our results are even more stronger in magnitude when the ETF management fees

and expenses are accounted for. The pre-fees annualized alpha for the difference portfolio of SRI ETFs and its passive benchmark is 8.28% in the final two years of our sample. This has important implications for investment planners as we show that in recent years, incorporating ESG preferences in ETF investment decisions could potentially help investors beat both the markets and the passive benchmark ETFs.

Since ETFs and SRI are both relatively new financial innovations and have both had significant growth after the 2008-2009 financial crisis, by studying SRI ETFs that combine these two innovations together in a single financial instrument we are also able to observe their trends. Indeed, the unique data sample that we present in this paper shows that majority of the SRI ETFs were issued after 2015. While there were only 24 SRI ETFs available for U.S. investors between 2010 to 2015, this number grew five times in the next 6 years. This might also explain the poor performance of SRI ETFs in the first few years of our sample period. As the SRI ETF industry developed and grew dramatically after 2015, more competition and greater investor demand together may have incentivized the fund managers to identify better screening strategies and offer superior products. As a result, the SRI ETFs could have improved in performance over the years. Our empirical tests on these two explanations for SRI ETFs' performance reveal that their declining market concentration may have most likely benefited SRI ETFs and fuelled their outperformance in recent years.

Furthermore, we provide insights on the strategies (Inclusion or Exclusion) and attributes (ESG and product-related) that drive the under-/over-performance of SRI ETFs over the benchmarks. Previous evidence points out that the different screening criteria and ESG dimensions may have a different effect on financial performance (Derwall et al., 2005; Kempf and Osthoff, 2007; Derwall et al., 2011). When we divide the SRI ETFs into different categories based on their screening style and ESG attributes, Inclusion strategy (positive screening) demonstrates significant abnormal returns over the years. The SRI ETFs that employ positive screens generate an alpha of -6.31% and 7.81% annually for the full sample (January 2010 to December 2020) and last two year (January 2019 to December 2020) periods respectively. In particular, Environmental Inclusion shows sig-

nificant negative annual alpha of 5.61% in the full sample period, and a huge positive abnormal returns of 13.80% when the sample is narrowed down to the last two years. Our results are consistent with the argument that using exclusion strategy may result in increased risk and lower returns (Humphrey and Tan, 2014).

To conclude, this paper shows that the financial performance and sustainable attributes need not be mutually exclusively when investing strategy involves ETFs. By offering a transparent and cost-efficient investment instruments, SRI or ESG ETFs have a huge potential to contribute to the cause of sustainability and sustainability development. As Nason (2020) states "[i]f you thought sustainable investing was just a do-gooder approach, it's time to take another look." For investors who want to integrate social responsibility and sustainability in their investment choice, we show that the financial underperformance may not always be a deterrent especially when they choose ETFs that incorporate Environmental Inclusion strategy. In other words, by choosing to invest in SRI ETFs that employ positive environmental screens, investors can "do good while doing well."

Appendix A. Alternative asset pricing models

We use alternative asset pricing models to check the robustness of all our results that employ the Carhart (1997) four-factor model presented in Section 3.4.1. In place of the four-factor model, we use a) the five-factor model (Fama and French, 2015), b) the three-factor model (Fama and French, 1993), and c) capital asset pricing model (CAPM).[8]

For the five-factor model, the following model is estimated:

$$R_{i,t} = \alpha + \beta_1 * RMRF_t + \beta_2 * SMB_t + \beta_3 * HML_t + \beta_4 * RMW_t + \beta_4 * CMA_t + \epsilon_{i,t}, \quad \text{(A.1)}$$

where $R_{i,t}$ is the excess return for portfolio i over the risk-free rate for each day t, and the five Fama-French factors are represented by $RMRF_t$ (market factor), SMB_t (size factor), HML_t (book-to-market or value factor), RMW_t (profitability factor), and CMA_t (investment factor).

In the three-factor model, we estimate:

$$R_{i,t} = \alpha + \beta_1 * RMRF_t + \beta_2 * SMB_t + \beta_3 * HML_{it} + \epsilon_{i,t}, \quad \text{(A.2)}$$

with the three Fama-French factors represented by three of the above five factors, i.e., $RMRF_t$ (market factor), SMB_t (size factor), and HML_t (book-to-market or value factor).

Despite the simplicity of CAPM, its mean-variance framework can provide meaningful insights (Ferson et al., 2013). Thus, lastly, we apply:

$$R_{i,t} = \alpha + \beta_1 * RMRF_t + \epsilon_{i,t}, \quad \text{(A.3)}$$

[8]We also considered using Pástor and Stambaugh (2003) liquidity factor, but omitted it from the analysis due to the lack of data availability for recent years.

Table A.I

Robustness check for Table III using Alternate Asset Pricing Models

This table reports alternative asset pricing model alphas for the sustainable SRI ETF portfolio, the benchmark S&P500 ETF portfolio, and the difference portfolio that tracks a long SRI ETFs short non-SRI S&P500 ETFs hedge. Alphas are estimated using equally-weighted portfolios of the daily ETF returns from 2010 to 2020. Three different time periods are shown: Full Period (January 2010 to December 2020), Half Period (January 2015 to December 2020), and Last 2 Years (January 2019 to December 2020). All alphas are annualized for ease of interpretation and expressed in percentage. Panel A reports the alphas from Fama and French (2015) five factor model, Panel B summarizes the alphas from Fama and French (1993) three factor model, and Panel C shows the alphas using CAPM. Significance levels are depicted by *** at 1%, ** at 5%, and * at 10%. The t-stats are reported in the brackets.

Panel A: Fama-French 5 Factors Model			
Time Periods:	Full Period	Half Period	Last 2 Years
SRI ETF Portfolio	-6.82***	0.64	4.91*
	[-3.27]	[0.28]	[1.57]
S&P500 ETF Portfolio	-2.07***	-2.38***	-2.56**
	[-3.27]	[-3.71]	[-2.55]
Difference Portfolio	-4.75***	1.49	7.37**
	[-2.92]	[0.60]	[2.14]
Panel B: Fama-French 3 Factors Model			
Time Periods:	Full Period	Half Period	Last 2 Years
SRI ETF Portfolio	-7.02***	0.52	4.84*
	[-3.34]	[0.23]	[1.55]
S&P500 ETF Portfolio	-1.85***	-2.28***	-2.40**
	[-2.87]	[-3.48]	[-2.24]
Difference Portfolio	-5.17***	1.74	7.24**
	[-3.01]	[0.54]	[2.32]
Panel C: Capital Asset Pricing Model (CAPM)			
Time Periods:	Full Period	Half Period	Last 2 Years
SRI ETF Portfolio	-7.90***	-0.52	5.19*
	[-3.61]	[-0.22]	[0.84]
S&P500 ETF Portfolio	-1.81**	-2.42***	-2.54**
	[-2.49]	[-3.06]	[-2.39]
Difference Portfolio	-6.09***	1.80	7.72**
	[-3.26]	[0.17]	[2.27]

Table A.II

Robustness check for Table IV using Alternate Asset Pricing Models

This table reports all the coefficient estimates for the sustainable SRI ETF portfolio, the benchmark S&P500 ETFs portfolio, and the difference portfolio that tracks a long SRI ETFs short non-SRI S&P500 ETFs hedge using alternative asset pricing models. The objective is to show that the results shown in Section 4 are not driven by the four-factor model applied in Table IV. We report the summary of estimations for the full sample period (January 2010 to December 2020) using Fama and French (2015) five factor model (Panel A), Fama and French (1993) three factor model (Panel B), and CAPM (Panel C). Similar to Table A.I, all alphas are annualized for ease of interpretation and expressed in percentage. Significance levels are shown using ***, **, and * at 1%, 5%, and 10% respectively. The t-stats for each of the coefficients are reported in the brackets.

Panel A: Fama-French 5 Factors Model							
	Alpha	*RMRF*	*SMB*	*HML*	*RMW*	*CMA*	*Adj.*R^2
SRI ETF Portfolio	-6.82***	0.95***	0.19***	0.09***	-0.13***	0.08**	0.933
	[-3.27]	[116.87]	[11.85]	[6.06]	[-5.53]	[2.53]	
S&P500 ETF Portfolio	-2.07***	0.97***	-0.12***	0.00***	0.07***	0.03***	0.992
	[-3.27]	[395.02]	[-24.59]	[-0.06]	[9.53]	[3.50]	
Difference Portfolio	-4.75***	0.95***	0.18***	0.08***	-0.14***	0.07**	0.928
	[-2.92]	[112.65]	[11.11]	[5.25]	[-5.79]	[2.24]	
Panel B: Fama-French 3 Factors Model							
	Alpha	*RMRF*	*SMB*	*HML*	*Adj.*R^2		
SRI ETF Portfolio	-7.02***	0.95***	0.21***	0.10***	0.932		
	[-3.34]	[122.61]	[13.54]	[8.21]			
S&P500 ETF Portfolio	-1.85***	0.96***	-0.13***	0.01***	0.992		
	[-2.87]	[404.52]	[-27.72]	[3.16]			
Difference Portfolio	-5.17***	0.95***	0.20***	0.09***	0.927		
	[-3.01]	[118.29]	[12.86]	[7.05]			
Panel A: Capital Asset Pricing Model (CAPM)							
	Alpha	*RMRF*	*Adj.*R^2				
SRI ETF Portfolio	-7.90***	0.99***	0.925				
	[-3.61]	[128.11]					
S&P500 ETF Portfolio	-1.81**	0.95***	0.990				
	[-2.49]	[369.67]					
Difference Portfolio	-6.09***	0.99***	0.921				
	[-3.26]	[124.14]					

References

Albuquerque, R., Y. Koskinen, S. Yang, and C. Zhang (2020). Resiliency of environmental and social stocks: An analysis of the exogenous COVID-19 market crash. *Review of Corporate Finance Studies 9*(3), 593–621.

Arnott, R. D. and M. Darnell (2003). Active versus passive management: Framing the decision. *Journal of Investing 12*(1), 31–36.

Barber, B. M. and T. Odean (2000). Trading is hazardous to your wealth: The common stock investment performance of individual investors. *Journal of Finance 55*(2), 773–806.

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

Bebchuk, L., A. Cohen, and A. Ferrell (2009). What matters in corporate governance? *Review of Financial Studies 22*(2), 783–827.

Bello, Z. Y. (2005). Socially responsible investing and portfolio diversification. *Journal of Financial Research 28*(1), 41–57.

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

Blitz, D. and F. J. Fabozzi (2017). Sin stocks revisited: Resolving the sin stock anomaly. *Journal of Portfolio Management 44*(1), 105–111.

Bodie, Z., A. Kane, and A. Marcus (2014). *EBOOK: Investments-Global edition.* McGraw Hill.

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

Box, T., R. L. Davis, and K. P. Fuller (2019). ETF competition and market quality. *Financial Management 48*(3), 873–916.

Brzeszczyński, J. and G. McIntosh (2014). Performance of portfolios composed of British SRI stocks. *Journal of Business Ethics 120*(3), 335–362.

Cao, S. S., G. Ma, J. W. Tucker, and C. Wan (2018). Technological peer pressure and product disclosure. *Accounting Review 93*(6), 95–126.

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

Carpenter, J. N. and A. W. Lynch (1999). Survivorship bias and attrition effects in measures of performance persistence. *Journal of Financial Economics 54*(3), 337–374.

Chakrabarty, B., S. B. Lee, and N. Singh (2017). Doing good while making money: Individual investor participation in socially responsible corporations. *Management Decision 55*(8), 1645–1659.

Chevalier, J. and G. Ellison (1999). Are some mutual fund managers better than others? Cross-sectional patterns in behavior and performance. *Journal of Finance 54*(3), 875–899.

Climent, F. and P. Soriano (2011). Green and good? The investment performance of US environmental mutual funds. *Journal of Business Ethics 103*(2), 275–287.

Cremers, M., A. Petajisto, and E. Zitzewitz (2012). Should benchmark indices have alpha? Revisiting performance evaluation. National Bureau of Economic Research Working Paper.

Delfeld, C. T. (2007). *ETF Investing Around the World: A Guide to Building a Global ETF Portfolio.* CIIR.

Dellva, W. L. (2001). Exchange-traded funds not for everyone. *Journal of Financial Planning 14*(4), 110.

Derwall, J., N. Guenster, R. Bauer, and K. Koedijk (2005). The eco-efficiency premium puzzle. *Financial Analysts Journal 61*(2), 51–63.

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

Ding, W., R. Levine, C. Lin, and W. Xie (2021). Corporate immunity to the COVID-19 pandemic. *Journal of Financial Economics 141*(2), 802–830.

Dumitrescu, A. and J. Gil-Bazo (2016). Information and investment under uncertainty. *Economics Letters 148*, 17–22.

Dumitrescu, A. and J. Gil-Bazo (2018). Market frictions, investor sophistication, and persistence in mutual fund performance. *Journal of Financial Markets 40*, 40–59.

Dumitrescu, A. and M. Zakriya (2021). Stakeholders and the stock price crash risk: What matters in corporate social performance? *Journal of Corporate Finance 67*, 101871.

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

Elton, E. J., M. J. Gruber, and A. de Souza (2019). Passive mutual funds and ETFs: Performance and comparison. *Journal of Banking & Finance 106*, 265–275.

Fabozzi, F. J. and H. M. Markowitz (2011). *The theory and practice of investment management: Asset allocation, valuation, portfolio construction, and strategies*, Volume 198. John Wiley & Sons.

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

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

Fama, E. F. and J. D. MacBeth (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy 81*(3), 607–636.

Fatemi, A., I. Fooladi, and H. Tehranian (2015). Valuation effects of corporate social responsibility. *Journal of Banking & Finance 59*, 182–192.

Ferson, W., S. Nallareddy, and B. Xie (2013). The "out-of-sample" performance of long run risk models. *Journal of Financial Economics 107*(3), 537–556.

French, K. R. (2008). Presidential address: The cost of active investing. *Journal of Finance 63*(4), 1537–1573.

Friedman, M. (1970). The social responsibility of business is to increase its profits. *New York Times Magazine*. 13th September.

Gil-Bazo, J. and P. Ruiz-Verdú (2009). The relation between price and performance in the mutual fund industry. *Journal of Finance 64*(5), 2153–2183.

Gil-Bazo, J., P. Ruiz-Verdú, and A. A. Santos (2010). The performance of socially responsible mutual funds: The role of fees and management companies. *Journal of Business Ethics 94*(2), 243–263.

Girard, E. C., H. Rahman, and B. A. Stone (2007). Socially responsible investments: Goody-two-shoes or bad to the bone? *Journal of Investing 16*(1), 96–110.

Gruber, M. J. (1996). Another puzzle: The growth in actively managed mutual funds. *Journal of Finance 51*(3), 783–810.

Guenster, N. (2012). *Performance Implications of SR Investing: Past versus Future*, Chapter 23, pp. 443–454. John Wiley & Sons, Ltd.

Guercio, D. D. and J. Reuter (2014). Mutual fund performance and the incentive to generate alpha. *Journal of Finance 69*(4), 1673–1704.

Hartzmark, S. M. and A. B. Sussman (2019). Do investors value sustainability? A natural experiment examining ranking and fund flows. *Journal of Finance 74*(6), 2789–2837.

Hong, H. and M. Kacperczyk (2009). The price of sin: The effects of social norms on markets. *Journal of Financial Economics 93*(1), 15–36.

Humphrey, J. E. and D. T. Tan (2014). Does it really hurt to be responsible? *Journal of Business Ethics 122*(3), 375–386.

Joliet, R. and Y. Titova (2018). Equity SRI funds vacillate between ethics and money: An analysis of the funds' stock holding decisions. *Journal of Banking & Finance 97*, 70–86.

Kacperczyk, M., S. Van Nieuwerburgh, and L. Veldkamp (2016). A rational theory of mutual funds' attention allocation. *Econometrica 84*(2), 571–626.

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

Kharma, C. and N. Eugster (2021). Is competition beneficial? The case of exchange traded funds. *International Review of Financial Analysis 76*, 101789.

Lettau, M. and A. Madhavan (2018). Exchange-traded funds 101 for economists. *Journal of Economic Perspectives 32*(1), 135–54.

Meziani, A. S. (2014). Investing with environmental, social, and governance issues in mind: From the back to the fore of style investing. *Journal of Investing 23*(3), 115–124.

Nason, D. (2020). Sustainable investing is surging. How to decide if it's right for you. *Forbes*. 5th November.

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

Omura, A., E. Roca, and M. Nakai (2021). Does responsible investing pay during economic downturns: Evidence from the COVID-19 pandemic. *Finance Research Letters 42*, 101914.

Pástor, L. and R. F. Stambaugh (2003). Liquidity risk and expected stock returns. *Journal of Political Economy 111*(3), 642–685.

Pavlova, I. and M. E. de Boyrie (2021). ESG ETFs and the COVID-19 stock market crash of 2020: Did clean funds fare better? *Finance Research Letters*, 102051.

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

Renneboog, L., J. Ter Horst, and C. Zhang (2008b). Socially responsible investments: Institutional aspects, performance, and investor behavior. *Journal of Banking & Finance 32*(9), 1723–1742.

Renneboog, L., J. Ter Horst, and C. Zhang (2011). Is ethical money financially smart? Nonfinancial attributes and money flows of socially responsible investment funds. *Journal of Financial Intermediation 20*(4), 562–588.

Riedl, A. and P. Smeets (2017). Why do investors hold socially responsible mutual funds? *Journal of Finance 72*(6), 2505–2550.

Rivoli, P. (1995). Ethical aspects of investor behavior. *Journal of Business Ethics 14*(4), 265–277.

Sabbaghi, O. (2011). The behavior of green exchange-traded funds. *Managerial Finance 37*(5), 426–441.

Schueth, S. (2003). Socially responsible investing in the United States. *Journal of Business Ethics 43*(3), 189–194.

Waddock, S. A., C. Bodwell, and S. B. Graves (2002). Responsibility: The new business imperative. *Academy of Management Perspectives 16*(2), 132–148.