Uncovering Green Foreign Direct Investment: A Novel LLM-Based Approach to Classifying Sustainable Investments[1]

Juan Alvarez-Vilanova

Riccardo Crescenzi

Lee Mager

London School of Economics & Political Science

FEBRUARY 2025

ABSTRACT

Foreign Direct Investment (FDI) is a crucial source of capital, technology and skills for the green transition, yet we lack a robust methodology for identifying 'green' FDI beyond investment in renewable energies. This paper develops a novel approach that leverages the newly published EU Taxonomy of Sustainable Finance and the rapidly advancing capabilities of pre-trained Large Language Models (LLM) to accurately, consistently and reliably identify green FDI projects in all sectors of the economy. This is, to our knowledge, the first application of LLMs specifically for classifying green investments. Our methodology reveals that 1 in 5 inward FDI USD into the European Union (EU) and the United Kingdom (UK) between 2013 and 2023 can be considered green, of which half lie beyond renewable energy and would otherwise not be fully picked up by existing approaches as capable of contributing to sustainable development. In terms of geography, our analysis unveils a new geography of green FDI that goes well beyond 'sunny' and 'windy' regions typically associated with renewable energies FDI. These descriptives demonstrate that, beyond acting as a source of finance for vital renewable energy infrastructure,

[1] This paper is part of "D3.2 - Report on the effectiveness of public policies for recovery and resilience in the era of pandemic and of digital and green transitions" produced in Work Package 3 – "The interconnections of in-market drivers of inequality with other processes and their impact across social groups and spatial scales" of the ESSPIN (Economic, Social and Spatial Inequalities in Europe in the Era of Global Mega-trends) Project. ESSPIN is a Horizon Europe Project supported by the European Commission under contract No.101061104. For more information on the project, its partners, and contributors please see ESSPIN website https://www.esspinhorizon.eu/

FDI in Europe is beginning to serve as a vehicle through which to decarbonise other sectors of the economy, including those (such as manufacturing and heavy industries) which are associated with high direct and indirect carbon emissions.

Keywords: Foreign Direct Investment, Green FDI, Sustainable Development, Green Transition, Large Language Models.

JEL Codes: F21; F23; O24; R58

EXTENDED ABSTRACT

Foreign Direct Investment (FDI) is a crucial source of capital, technology and skills for the green transition (UNCTAD, 2023; Rogelji et al, 2018), which is in turn essential for averting the worst effects of climate change and meeting the goals of the 2015 Paris Agreement. However, we lack a comprehensive classification system for identifying what constitutes green FDI.

Up until very recently, efforts to do this have been hampered by two key factors. Firstly, the lack of a universal and comprehensive definition of what makes FDI green. Previous classifications have often been limited to one-dimensional categorisations, based either on identifying FDI within certain 'green' sectors or on calculating Greenhouse Gas (GHG) emissions. Yet the publication in 2023 of the full EU Taxonomy of Sustainable Finance (EUTSF), which offers a comprehensive framework for classifying green activities across all sectors of the economy, changes this.

Secondly, precisely identifying green FDI across all sectors of the economy requires a bottom-up approach that evaluates individual investments for the presence of green activities, and is therefore extremely time-intensive, even with a comprehensive framework like the EUTSA to guide the decision. Here, recent advancements in the abilities of pre-trained Large Language Models (LLMs), are a game-changer. In the past 5 years, LLMs have significantly enhanced their capabilities in performing complex Natural Language Processing (NLP) tasks. GPT-4, in particular, has shown a promising ability to address intricate classification challenges, with the potential to streamline and improve the accuracy of labelling and reasoning even for abstract concepts (Bubeck, Chandrasekaran et al., 2023).

Against this backdrop, we propose a novel approach that leverages the newly published EUTSF and the rapidly advancing capabilities of pre-trained LLMs to robustly identify green FDI projects. By integrating the EU Taxonomy criteria directly within our model prompts, we enable GPT-4 to evaluate and determine which investments qualify as green based on detailed project descriptions and the relevant EUTSA criteria.

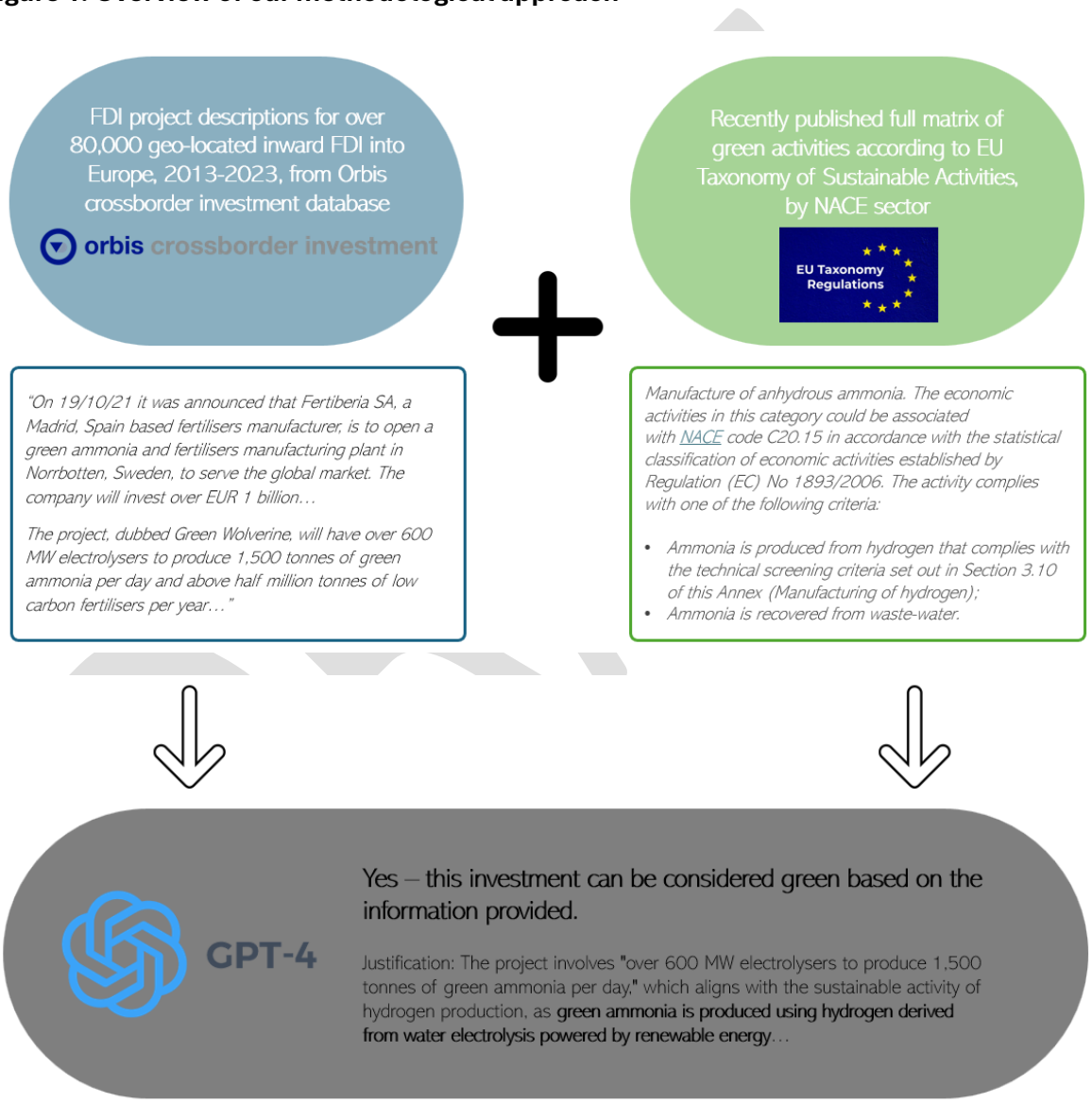
We implement this on a dataset of all inward FDI into Europe from 2013-2023, finding that 1 in every 5 FDI USD can be classified as green. Over half of such green investments lie beyond the 'core' green sectors, typically defined in the literature as renewable energy and waste management , and would otherwise not be picked up by less nuanced, sector-based approaches to labelling green FDI. Moreover, we demonstrate that the decisions made by GPT-4 are accurate in the sense of aligning with the conclusions of human evaluators, and valid in the sense of making decisions as instructed – by considering exclusively the EUTSA criteria and investment information provided, rather than parametrized knowledge on the subject.

Our work contributes to a rapidly-expanding literature on the application of NLP techniques in sustainable finance research. LLMs like GPT-4 have been used to assess corporate climate transition disclosures (Colesanti-Senni et al., 2024) and to draw on sources like the Inter-Governmental Panel on Climate Change (IPCC) reports to answer complex queries pertaining to climate change (Ashraf-Vaghefi et al., 2023). While recent studies like those by Jain et al. (2024) have used LLMs to estimate GHG emissions of various investments, ours is the first, to our knowledge, to apply LLMs specifically for classifying green investments according to a comprehensive framework like the EUTSA.

Our methodology

Our methodology for identifying 'green' Foreign Direct Investment (FDI) projects is developed and tested on a dataset of inward FDI investments made in European countries (EU27+UK) between 2013 and 2023. This dataset, sourced from Orbis-Crossborder (by Bureau van Dijk/Moody's Analytics), provides detailed descriptions for the majority of FDI projects, including information about the investor, investment location, capital value, job creation, and a textual description of the investment (including, often, its rationale).

Figure 1: Overview of our methodological approach



To identify which of these investments are green, we employ the GPT-4 LLM to precisely and consistently determine whether each investment aligns with the activities listed in the EU Taxonomy of Sustainable Activities (EUTSF). This is done by providing the LLM agent with two key pieces of information: firstly, the textual description of each investment in our dataset; secondly, we extract and provide the text on the activities considered 'green' in that particular NACE sector

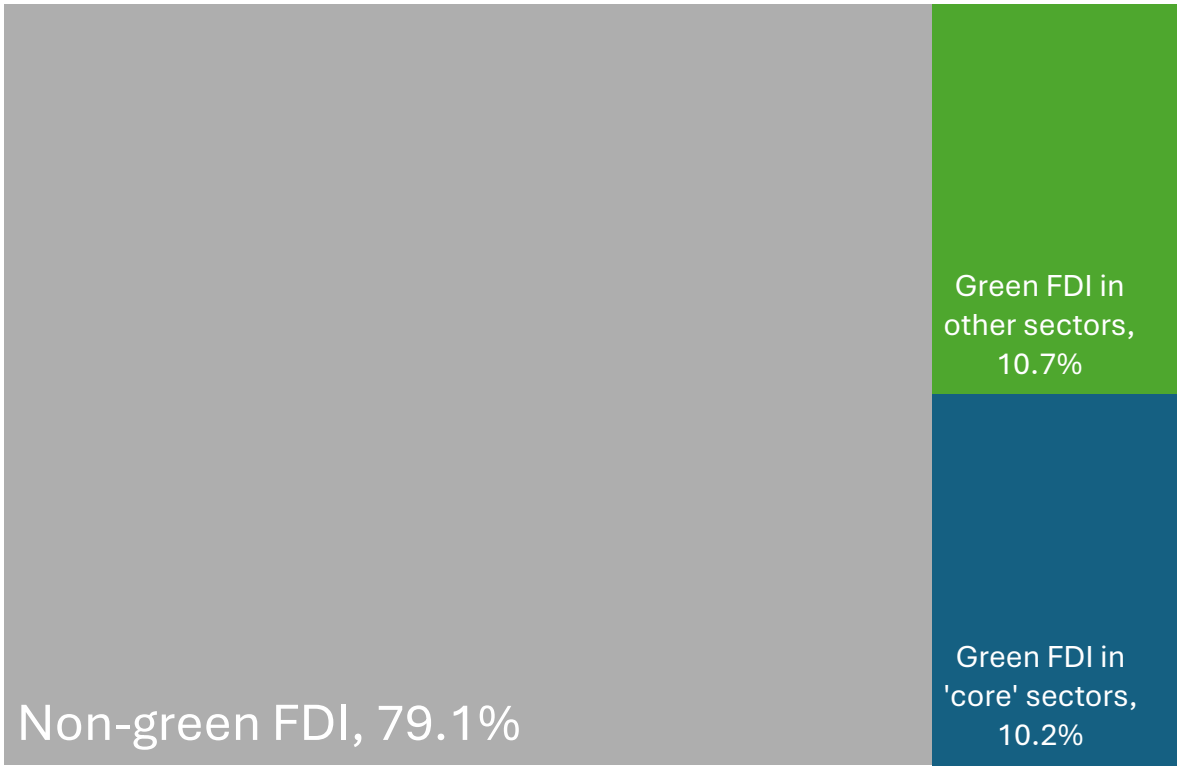
as per the EUTSA. These two pieces of information are combined into tailored prompts for each inward FDI project.

The prompt is designed to instruct the model to base its decision exclusively on the provided information, rather than inferring anything about the investment or sourcing related but irrelevant information from its own corpus. We demonstrate that the decisions made by GPT-4 are accurate in the sense of aligning with the conclusions of human evaluators. In the in-person trials organised to gauge the LLM's accuracy, human evaluators agreed with the outcome of GPT-4 in 83%-93% of cases. It is also valid in the sense of making decisions as instructed – by considering exclusively the EUTSF criteria and investment information provided, rather than parametrized knowledge on the subject. Notably, our methodology has also demonstrated reproducibility across different model parameters and prompts, underscoring the inherent capability of LLMs to consistently perform such complex classifications despite the stochastic nature of text generation in models like GPT-4.

Preliminary results

After processing each FDI project through the methodology described above, we find that just over 1 in every 5 USD of inward FDI into Europe between 2013-2023 (inclusive) can be considered green. Interestingly, we find that just over half of this (10.7% of all inward FDI) is inward FDI that falls outside of the 'core' green sectors as defined in section 2.2 (see Figure 2, below). In other words, this constitutes the share of FDI that may not be picked up as 'green' through coarser, sector-based approaches to identifying green FDI that have previously been employed in the empirical literature.

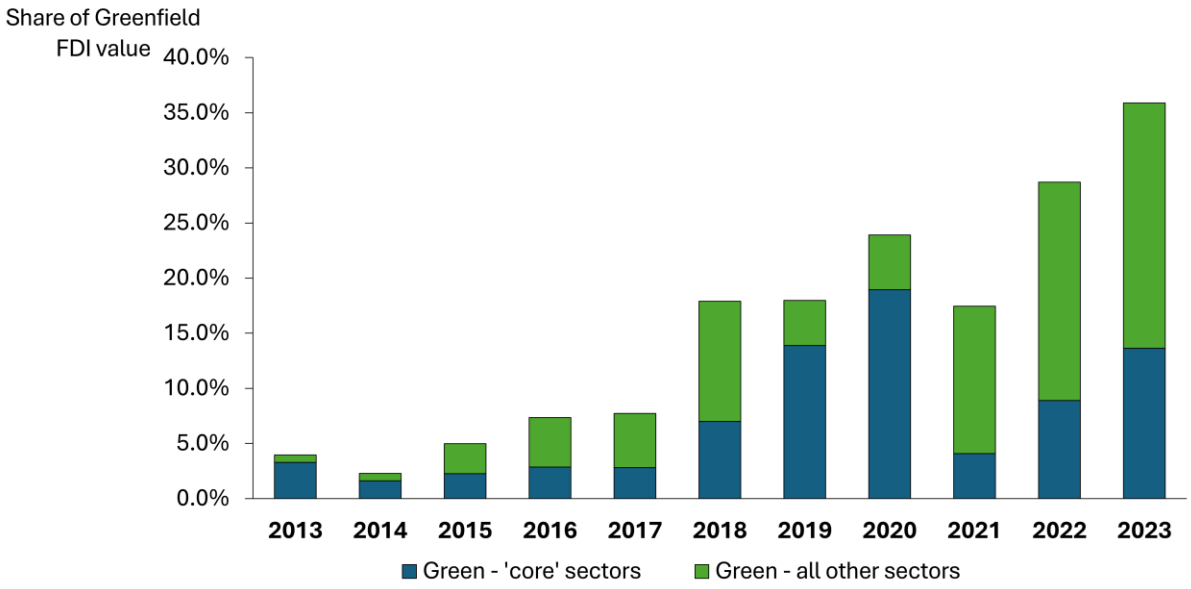
Figure 2: Share of inward FDI into EU27+UK that is 'green' (in terms of FDI value, 2013-2023)



Source: Own elaboration with data from Orbis-crossborder / Bureau van Dijk (2023)

Moreover, as can be seen in Figure 3 below, the share of green FDI has consistently grown over the past 11 years. In 2022 and 2023, in particular, over 25% of all inward FDI into the EU27+UK was green, with green FDI from outside of the core sectors making up the majority. This aligns with the limited existing analysis that seeks to ascertain green FDI trends, namely recent trends reported by UNCTAD (2023) and BCG (2022) which find that certain, non-energy sectors like manufacturing increasingly attract FDI into specific activities, like EV manufacturing, which could be considered green 13. This suggests that, beyond acting as a source of finance for vital renewable energy infrastructure, FDI in Europe is beginning to serve as a vehicle through which to decarbonise other sectors of the economy, including those (such as manufacturing and heavy industries) which are associated with high direct and indirect carbon emissions (Hertwich & Wood, 2018; OECD, 2023).

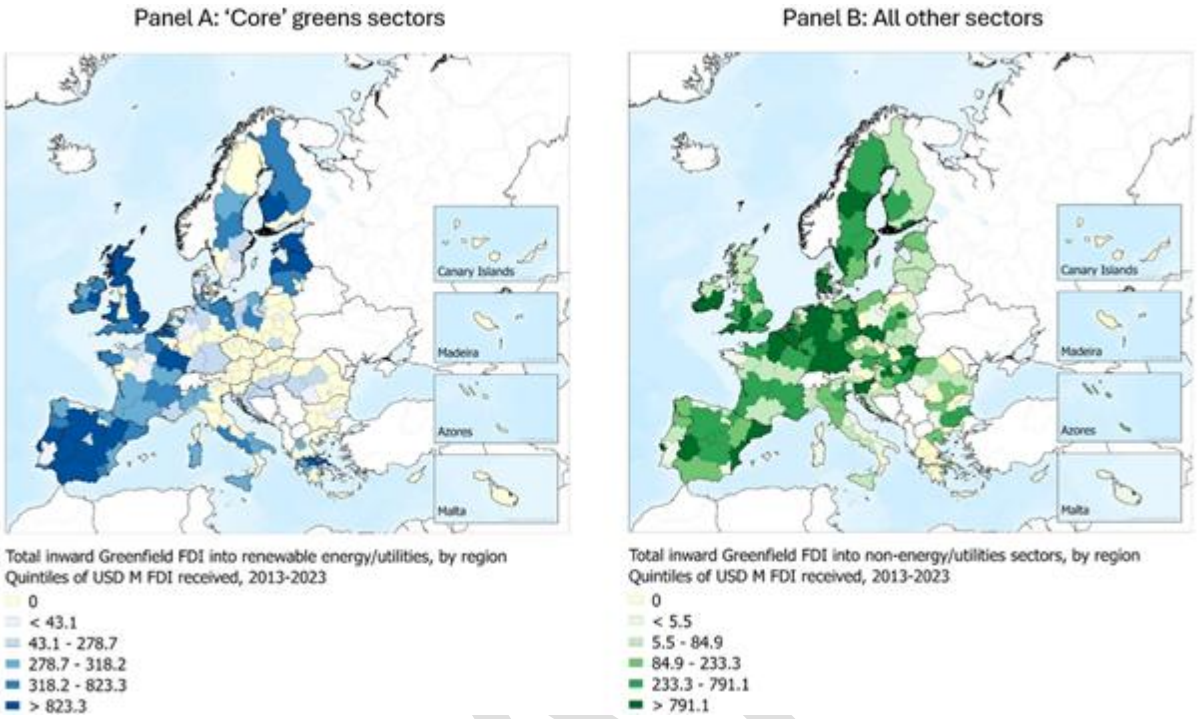
Figure 3: Evolution in the share of inward FDI into EU27+UK that is 'green' (in terms of FDI value, 2013-2023)



Source: Own elaboration with data from Orbis-crossborder / Bureau van Dijk (2023)

Figure 3, above, shows the regional distribution of both types of green FDI into the EU27 plus the UK. It can seen that green FDI in the energy sector (Panel A) predominantly clusters in sunny regions (central Spain and Portugal, plus parts of Southern France and Southern Greece) and windy regions (Eastern England, Scotland, plus several regions around the Irish and Baltic seas) where conditions for solar and offshore wind power generation (respectively) are known to be best. By contrast, the distribution of green FDI beyond the energy sector (Panel B) paints a more varied picture. In this case, the top performing regions include both heavily industrial areas of Eastern Europe (particularly in Hungary, Slovakia and Poland), which have attracted a high level of FDI into 'green manufacturing' activities such as the production of EVs and/or EV batteries, plus a number of advanced metropolitan regions in Western Europe such as Catalonia, Ile-de-France, the English West Midlands or Baden-Wurttemberg, where green inward FDI constitute a mix of heavy (and traditionally high-polluting) manufactures with investments in green R&D, technical/professional services and IT services.

Figure 4: The geographical distribution of green FDI in EU2+UK: 'core' green FDI vs green FDI in all other sectors



Finally, Figure 5 below, contains two such examples of green FDI projects that falls outside of the 'core' renewable energy sector: an investment by Spanish MNE Fertiberia to build a green ammonia plant in Norrbotten, Sweden, in 2021, and an investment by Swiss MNE ABB to open an e-mobility R&D lab in Delft, the Netherlands. Such investments would not have been picked up and labelled as green under traditional, sector-based approaches to defining the green economy, however by applying a classification methodology based on cross-cutting green activities as defined by the EUTSA, we are able to identify investments like this which make a vital contribution to the green transition.

Figure 5. Four examples of a green FDI projects identified: two that fall within the 'core' green sectors (left), and two that are outside of the 'core' green sectors

A $295mn investment by MNE in the "Electricity generation" sector (NACE D35):

"On 01/12/21 it was announced that Amazon.com inc, an online goods retailer platform provider, is to open ***a wind power plant*** *in Scotland, UK"*

A $1.1bn investment by MNE in the "Manufacturing – chemicals" sector (NACE C20):

"On 19/10/21 it was announced that Fertiberia SA, a Madrid, Spain based fertilisers manufacturer, is to open ***a green ammonia and fertilisers manufacturing plant*** *in Norrbotten, Sweden, to serve the global market."*

A $5.3mn investment by MNE in the "Waste collection and treatment" sector (NACE E38):

"On 28/02/23 it was announced that Premier Tech Water and Environment SL, an environmental research services provider, is to open ***a water and wastewater treatment systems manufacturing plant*** *in Amposta, Spain"*

A $10mn investment by MNE in the "Scientific R&D" sector (NACE M72):

"On 09/12/20 it was announced that ABB ASEA Brown Boveri Ltd (ABB), a Zurich, Switzerland based electrical equipment manufacturer, ***has opened an e-mobility research and development centre*** *in Delft, Netherlands"*

Source: Own elaboration using data from Orbis-crossborder / Bureau van Dijk (2023)