ERASMUS UNIVERSITY ROTTERDAM

ERASMUS SCHOOL OF ECONOMICS

MSc Economics & Business

Specialization Financial Economics

Modelling the Carbon Footprint of Blockchains

Author:	Andre Carlos Camargo Marcatto
Student number:	488782
Thesis supervisor:	Dr. Jan Lemmen
Second reader:	Dr. Jens Kvaerner
Finish date:	09/07/2023

PREFACE AND ACKNOWLEDGEMENTS

I want to acknowledge the support from Erasmus University, its board, and my Professor Jan Lemmen. Thank you Jan for providing relevant sources and constructive feedback, being empathetic and motivating me during my personal challenges.

I would also like to thank my family and partner for being by my side while I wrote my Master thesis.

Learning more about the intricacies of blockchains while applying the knowledge I acquired during my studies was a pleasure. I will remember this part of my young adulthood with joy.

ABSTRACT

This study explores the sustainability aspect of blockchain technology by focusing on its carbon footprint and externalities. With sustainability being a global priority, numerous countries have implemented regulations to promote a more sustainable future, while international agreements like the Paris Agreement aim for carbon neutrality. As a prominent distributed ledger technology (DLT), blockchain has gained attention due to its negative environmental externalities. The energy consumption required for processing blockchain transactions, particularly in the case of Bitcoin, has raised concerns. However, alternative DLTs, like Ethereum, have demonstrated the potential to reduce carbon emissions significantly. This research investigates the main drivers of the blockchain ecosystem's carbon footprint and explores the effects of structural breaks, protocol changes, and legislation. Findings highlight energy consumption as the primary driver, with additional variables such as hashing power and market prices found to be relevant. The results show dual Granger causality and cointegration within the variables. In one case, the ecosystem carbon footprint Granger causes effects in individual DLTs. In the long run, the variables have a negative long-term coefficient against an individual DLT. The DLTs have a predominant positive long-term relationship when considering the whole ecosystem. The study recommends regulations that permanently affect the relevant variables, which will also affect their long-term trend. The study also contributes to increasing awareness and understanding of blockchain's environmental impact, aiding in developing strategies for managing carbon emissions in DLTs.

Keywords: Blockchain, Sustainability, Carbon Footprint.

JEL Classification: C58, Financial Econometrics; C32, Time-Series Models; C36, Instrumental Variables Estimation; C51, Model Construction and Estimation; Q56, Sustainability; Q55, Technological Innovation; Q58, Government Policy

TABLE OF CONTENTS

List of Tables

List of Figures

CHAPTER 1 Introduction

Sustainability is one of the top priorities for the world. Most countries have signed international treaties or are developing regulations to encourage a more sustainable future. For example, 175 nations pledged to become carbon neutral by 2050 in the Paris Agreement, and China vowed to do the same by 2060 (UN, 2023 & Hook, 2020). Furthermore, from 2022, new regulations require companies with more than 500 employees in Europe to report standardised metrics of their current CO_2 emissions (UN, 2023). The regulation increases transparency and avoids greenwashing. Large corporations face increased regulation and accountability in terms of environmental standards. Small enterprises, on the other hand, face fewer similar trends.

One emerging technology that is raising alarms in academia and hitting newspapers' first pages due to its negative environmental externalities is Blockchain. In essence, a Blockchain is a Distributed Ledger Technology (DLT). A DLT is a method of storing valuable information online or offline. DLTs make valuable information accessible (encrypted) to everyone and in multiple copies. There are different types of tokens to store information, which depends on their purpose. Most multinational enterprises and countries are developing and implementing processes using DLTs as a foundation. Forbes gathers an overview of the most relevant projects annually with the "Block 50" article (2022). Companies are experimenting with new ways to keep records, increasing traceability and accountability between parties. Also, the interaction between multinationals with Blockchain projects, such as executing transactions and supply chain management, is being tested.

Estimations from Cambridge University (2023) show that the energy required to process one transaction on the most famous blockchain – Bitcoin – equals the energy needed to maintain an average American household for 30 days. Bitcoin's carbon footprint is estimated to be 74.31 Mt of CO_2 per year, which is comparable to Colombia's carbon footprint in 2021 (Cambridge, 2021). Bitcoin operations have far-reaching impacts.

Not all blockchains operate in the same way as Bitcoin. There exist alternative DLTs that, while sacrificing decentralisation, can significantly reduce their carbon emissions. This is done by minimising the computational resources required to sustain DLTs. For example, the Ethereum Blockchain (the second biggest by value) changed a core process: it decreased the carbon footprint by more than 99% and maintained its energy consumption relatively constant (Sarkar, 2022). I will elaborate specifically on Ethereum and Bitcoin operations and use that as a proxy for the DLT ecosystem.

Because energy expenditure for transactions is far from optimal, the viability of DLTs can be questioned, especially with DLT increasing its functionalities and exposure to individuals and institutions. Throughout 2019-2022, environmental concerns regarding blockchain technologies caused some pioneering nations to enforce restrictive legislation. For instance, China and eight other countries banned crypto mining (Quiroz Gutierrez, 2022). Moreover, a growing concern surrounds e-waste generated through maintaining the blockchains. Governments and academia are actively researching recent blockchain-related processes and their externalities from diverse perspectives.

The United Kingdom, which is often a forerunner in the regulation of financial markets, in February 2023 released a report stating that the environmental impact of cryptocurrencies is considerable. Further, the UK wants to create a framework to calculate ESG scores for cryptocurrencies (HM Treasury, 2023). That score could inform individuals investing in crypto assets about environmental risks. To do so, the British government is researching blockchains' carbon emissions and energy usage drivers (HM Treasury, 2023). There is also the perspective that cryptocurrencies can affect the electricity market (Fitch, 2022). This means that crypto operations put electricity prices under pressure, potentially driving them up. The goal of this research is to answer the following question:

"What are the main drivers of the carbon footprint from the blockchain ecosystem and its externalities between 2019-2022?"

By answering the research question, I provide insights into the primary causes of the carbon footprint from the blockchain ecosystem and peer-review the approaches that different researchers have taken while looking for structural breaks in the energy consumption of different blockchains. I could assist future developments in DLTs by increasing awareness of the leading causes of carbon emissions. Thus, different parties can have a more critical understanding of the carbon emission from their DTLs and could be able to manage carbon emissions more adequately.

While there are many findings, energy consumption in DLTs is the primary driver of the carbon footprint. Further, there is a structural break in the carbon footprint of Ethereum at a 5% confidence level. Bitcoin also has a relevant structural break in its carbon footprint at a 10% confidence level. After considering the structural breaks, other variables were also found relevant but not for the whole data sample. Those variables are hashing power and market prices. Collinearity was found while inspecting for Granger causality, and no relevant

Instrument variable was found. On the other hand, cointegration between the variables was found. Thus, their long-term relationship was modelled. The model found that energy, hashing power and price have a negative long-term relationship with individual DLTs. Lastly, the individual DLTs' carbon footprint has, in most cases, a positive long-term relationship with the ecosystem.

I discuss all the findings in relation to the literature and my expectations in Chapter 5. Before that, I provide an overview of background information, contextualising the research (chapter 2). With the literature at hand, I determined key factors and approaches to model carbon footprints from blockchains and enumerated several hypotheses. I determine the data in the third chapter and comment on its properties. The fourth chapter discusses the specific methods by which the research and analyses were conducted. The fifth chapter displays the results generated from the data analysis and discusses its intricacies. The last section will respond to the hypotheses, answer the research question, and discuss the nuances of my findings.

CHAPTER 2 Literature Review

In this chapter, I will touch upon the existing literature to explain how blockchains relate to sustainability, what blockchains' functionalities are, and consequently, derive four hypotheses.

2.1 Sustainability

Blockchains are constantly in the news due to their significant carbon footprint. For example, The Guardian (2022) writes about how crypto mining in the US contributes to raising utility bills. This news came after lawmakers in the US requested information from the country's five biggest crypto mining companies. The report states that the mining operations generate 1.6mln tons of CO_2 annually. In comparison, the average household in the EU emits around 10 tons of CO_2 annually (Zerofy, 2022). The Guardian interrogated the Rochester Institute of Technology about the subject. They believe that US citizens cope with higher energy bills to compensate the mining companies in the region for maintaining blockchains. In other words, crypto mining companies have energy-intensive operations, which put pressure on the energy prices in the region. As a result, individuals in the region see an increase in their utility bills. Mining companies are willing to maintain blockchains because they are able to generate profits by doing so. The regional government is expected to create legislation to protect individuals from absorbing the external cost that mining companies produce. Because, as it is now, the price increase generated by crypto-mining companies is passed to the average person. This is made by electricity companies increasing the price of the megawatt-hours for everyone in the region.

Numerous studies and reports have compared blockchain's carbon emissions to other countries. Bitcoin is most frequently used to refer to the entire blockchain ecosystem. This is the case because Bitcoin has the most users. Cambridge University wrote the best-known report on Bitcoin's carbon emissions. According to their calculations, Bitcoin alone has a similar yearly carbon footprint to Botswana, 52.1 and 52.3 megatons of C02, respectively (Cambridge University, 2023). But it is not the most precise and reliable comparison to put emissions into perspective. After all, blockchains do not run like nations.

Bitcoins have also been compared to gold. Bitcoin mining sets off 48.3 megatons of CO_2 less than gold per year (Cambridge University, 2023). This comparison relies on the idea that both bitcoins and gold can maintain their value over time and have a finite amount. Thus, both currencies are connected to the concept of scarcity. Another similarity is that both are hard to be destroyed. Such an idea is shared within major investment banks such as Goldman Sachs (Economic Times, 2022).

The average person uses Bitcoin to make transactions. The emission for a bitcoin transaction can be compared against the average CO_2 emission of a credit card such as a Visa. According to Best, one bitcoin transaction emits the CO_2 equivalent to 706,765 Visa transactions (2023). This comparison is not ultimately straightforward. Only internal Visa operations directly related to the execution of a transaction are considered when calculating the CO_2 emissions from a Visa transaction. Part of Visa's operations is outsourced to the international banking system (e.g., SWIFT), whereas Bitcoin operates with a completely integrated value chain. Thus, in this instance, the total cost of a transaction, considering the entire banking system, should be included in the calculation. Measuring and comparing transactions in bitcoins to other transactions is necessary for tracking sustainability, nevertheless challenging because of its large scale.

Beyond the features listed above, blockchains have additional capabilities. Because of these features, multi-national organisations like Unilever and federal authorities like the Swedish government are beginning to test blockchain in their operations. Because organisations and authorities are using the same technology as Bitcoin, similar carbon footprints can be expected from their blockchains if they are not constructed carefully. Therefore, it is crucial to comprehend what drives blockchains' carbon emissions. The next section expands on blockchain technology and its implementations.

2.2 What is blockchain?

Blockchain technology, on a superficial level, can be described as a manner to store value. The idea of value can be anything a person deems worth it, from a written note to land ownership. The information is stored in a Distributed Ledger Technology (DLT). The idea behind a distributed ledger is to store the same data in various locations while maintaining its integrity. This is made possible by computer software. Thus, a blockchain is a DLT. Traditionally, a ledger only keeps track of past transactions rather than necessarily the full details of each one (Merriam-Webster, 2023). For the vast majority of blockchains, this is also the case (Wang, 2021).

A blockchain allows individuals to prove ownership reliably. This mechanism was formally introduced in the white paper from Nakamoto (2008), where he/she explains how computers can activate a consensus in the ledger using logic, encryption, and competition elements. This consensus mechanism is called Proof of Work (PoW). Blockchains rely on consensus mechanisms to operate. A consensus mechanism is a set of guidelines or agreements that allows every node in a blockchain network to verify the legitimacy of all transactions (Lashkari & Musilek, 2021). There is another prominent consensus mechanism recently adopted by the second largest public blockchain Ethereum, which is Proof of Stake (PoS). This mechanism was mainly adopted with the goal of decreasing Ethereum's carbon footprint (Ethereum, 2023). In the next two sections, PoW and PoS are respectively explained in depth.

Oxford developed a framework to break down the fundamental of a blockchain (2022). That framework is named Oxford Blockchain Strategy Framework (OBSF). In such a framework, the DLT can be divided into three layers: protocol, network and application. The protocol layer consists of the basic operating principles, such as design expectations for speed, programmability and functionality, or even if the code is open or closed source. Now the network layer focuses on the infrastructure of the system. This entails setting the requirements for who can run a node, which nodes have reading or writing access, how difficult it is to calculate a nonce, and data storage requirements regulations. Lastly, the application layer focuses on the interaction between specific user cases develop to use the DLT system, such as trading native tokens. All blockchains have different approaches for their layers, thus allowing for diverse approaches to reaching a similar goal, which can result in distinctive externalities (Oxford, 2022).

OBSF also defined six principles that all DLTs must comply with; those principles are also benefits that are implicit in the system. The principles are automation, continuum, stakeholders, reconciliation, value transfer, and immutability. Automation refers to a predictable and repeatable process, continuum relates to a process that keeps on going, stakeholders implies that it accommodates multiple stakeholders in the value chain, and reconciliation means that there must be an agreement on the validity of the information generated by the process, value transfer conveys the element of value transferer, which not necessarily is monetary, and immutability expresses the idea that the information cannot be changed (Oxford, 2022).

As its name insinuates, a blockchain stores information in blocks and connects them by considering historical information. Each block has four components: the information from the previous block, the information from its own block, a time stamp, and a nonce. The nonce is a mathematical problem that relates to the information stored in the block. When a nonce is calculated, a hash is produced. The hashing rate of a cryptocurrency network can be used to assess its stability and security because hashing adds an encryption layer that protects the information against tampering. Because the nonce relies on the speed of the machines being used and/or the number of miners in a network, hashing rate representations might vary from network to network and even from miner to miner (Alam et al., 2021). For instance, the hashing rate for Bitcoin is measured in exahashes per second (EH/s) and is generated using the SHA-256 cryptographic algorithm (Singh, 2022). There are one quintillion hashes in an exahash. On the other hand, the current unit of measure for Ethereum speed is terahashes per second (TH/s). In one terahash, there are one trillion hashes (Ethereum, 2023). The TH/s will be utilized in order to improve blockchain comparison. In Figure 1, a schematization of a blockchain ledger structure is shown.

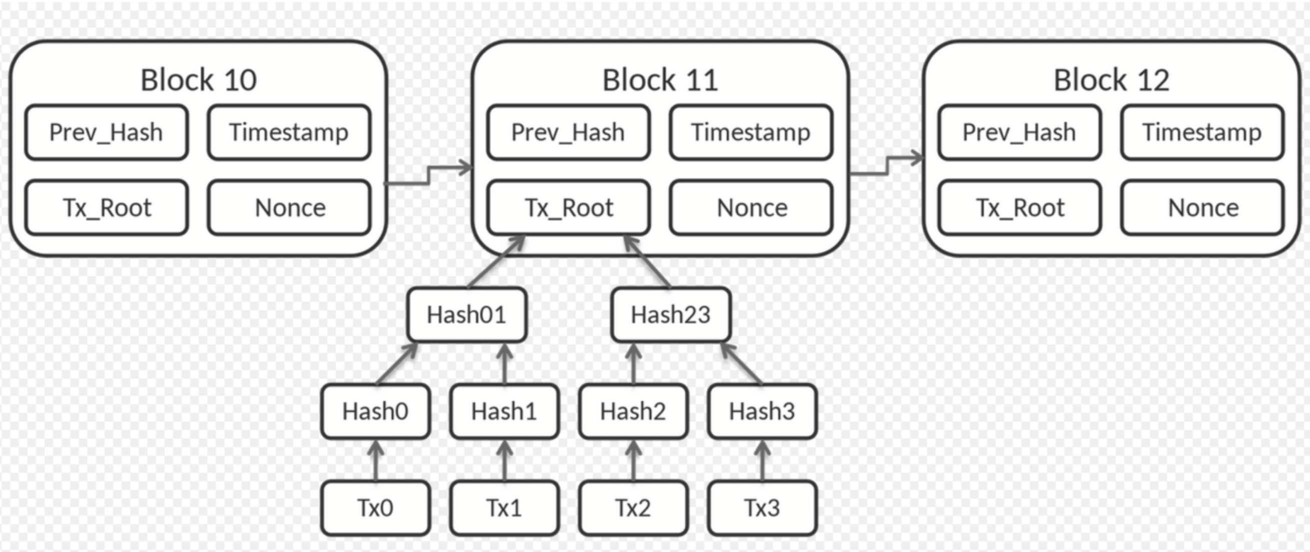

Figure 1 Blockchain structure

Adapted source: Lecture 2 Financial Economics Masters (Blockchain fundamental) 2022.

In a blockchain, the servers are referred to as “nodes.” Transactions are processed by nodes. The ledger can be maintained and expanded by certain nodes by adding blocks of transactions to the chain. These nodes are known as “miners” in (PoW) networks like Bitcoin due to their nature and similarities with gold. Most commonly miners are rewarded by receiving the native token from the blockchains, as receiving bitcoins by mining in the Bitcoin blockchain.

2.2.1 Proof of Work

PoW can be seen as a small competition where all the mining nodes attempt solving the nonce to successfully create a new block. Thus, the node that solves the nonce first gains the rights over the mining prize, which changes according to blockchain protocols. This competition has a significant impact on the computational power required to mine a new block (Gulli, 2020). This high computational power requirement also implies a high energy demand**.**

In a 1993 study, Cynthia Dwork and Moni Naor initially proposed the concept of employing computational work as a way to prevent spam and denial-of-service exploits. The term "proof of work" is considered to have appeared first in a 1999 essay by computer scientists Ari Juels and Markus Jakobsson. Their idea has been shaped by the computing community since the early 1990s. Just with Nakamoto's white paper, PoW was conceptualized to verify transaction between peers.

De Vries et al. (2022) speculates that Bitcoin has been operational for more than a decade without experiencing a significant outage or hack, which could be the most crucial evidence that PoW offers a greater level of security than alternative methods of consensus. While proof of work does provide the highest possible level of security and decentralization, it comes at a significant cost: it uses a significant quantity of energy. PoW is also used by many other cryptocurrencies that are not based on Bitcoin protocols, such as, Ethereum Classic, Monero, Zcash, Kadena, Ravencoin, Siacoin, and Horizen.

2.2.2 Proof of Stake

As was already mentioned, a different kind of consensus mechanism that has a smaller carbon footprint is becoming more and more popular (Pennella, 2023). This approach is called PoS. In a PoS system, network nodes place "stakes" of tokens for a certain amount of time in exchange for a chance to create the next block of transactions. The selected node, known as the "validator," will be given the block rewards in the form of the network's native token. But this comes at a cost, a loss in decentralization.

PoS requires validator nodes to lock their tokens in order to stake them. This means that one gives up the right to trade the token in the market while running the validator node. To compensate for the loss in liquidity, the validator node is rewarded once it's selected to mine a new node. There are also fines if a validator is not active when selected to check the information. Further, validator nodes are also implemented to double-check the information in the new block. PoS takes out the competition element of the consensus by adding a random factor that gives turns for the miners to verify. Since the turns are randomly chosen, validators and miners stay idle to perform their tasks when necessary instead of competing over the nonce. Activities per node demand less energy to maintain the system's operation. This allows for an overall decrease in the required computational power. PoS systems should have a somewhat constant energy usage.

PoS can solve some of the main issues PoW has, including energy consumption (because PoS uses less energy than PoW), transaction throughput (since PoS networks can handle more transactions than PoW networks); and scalability (because PoS networks scale more readily than PoW networks).

2.3 Carbon footprint

The energy consumption generated by maintaining a computer running a miner node does not incorporate all the factors that explain CO2 in a DLT system. The core problem is not the energy consumption by itself, but the carbon generated by its operations. There is a growing debate on the relevance of some variables, such as the price of native tokens, but there is an agreement on the main driver, which is the CO2 generated by the mining nodes (Ghosh & Bouri, 2022).

Different parties have estimated the carbon footprints of diverse DLTs systems. Most of them use the two largest open blockchains as a proxy for the ecosystem. This is acceptable because Bitcoin and Ethereum combined account for more than 80% of the world's cryptocurrency assets' power usage (The White House, 2022). In this research, I will also employ the summation of both blockchains as a proxy for the ecosystem. The carbon footprint is typically assessed based on the volume of transactions conducted over a period. However, it is important to note that this measure assumes a direct correlation between the number of transactions and the carbon footprint, which is not necessarily true as the blockchain's protocol defines the frequency of transactions over time. Nonetheless, utilising this measure will provide insights into how scalability can impact the carbon footprint. Hence, my approach will involve evaluating the carbon footprint of each blockchain on a per-transaction basis. Also, it is important to highlight that the carbon footprint of the DLTs is an estimation. Cambridge Center provides the most accepted estimation for DLTs' carbon footprint and energy consumption for alternative finance and Digiconomist. Because the carbon footprint of a blockchain is an estimation, it is also relevant to check if they are statistically different within them. I expect to find different carbon footprints per transaction due to their fundamental dissimilarities in its layers as defined by the OBSF. Thus, my first hypothesis is to test such expectations.

"*There is a statistical difference between the average carbon emission per transaction across different DLTs*. "

Testing for such a statistical difference will also allow for a better understanding of the underlying drivers of carbon footprints. This can be done by comparing the differences between the protocols, such as the maximum number of transactions per block or the type of consensus taken. In previous research, a relevant difference between PoW and PoS was found (de Vries et al., 2022). But no difference within DLTs with the same consensus mechanism was investigated. In this research, I will this occasion, by comparing Bitcoin and Ethereum's carbon footprint while they both followed the same consensus approach. The difference will be inspected both in absolute terms and in percentage changes. Possible differences in the carbon footprints of the DLTs, may emerge from the dissimilarities in their layers.

2.3.1 Relevant variables

The main cause of a carbon footprint comes from the computers trying to mine a new block (Gallersdörfer et al., 2020). The most appropriate way to derive the carbon footprint of a blockchain includes furthermore measuring how much energy was used to maintain the ecosystem and its external costs (Gallersdörfer et al., 2020). Measuring the energy used to maintain the computer running is a reasonable approach to estimating the carbon footprint of a blockchain.

A DLT system's energy consumption has historically been estimated using two major methods. One method is to gauge the consumption of a participant node that serves as a representative sample before extrapolating from that data. Another method for measuring energy consumption that brings valuable insights comes from UCL Centre of Blockchain Technologies (Platt et al., 2021). Their mathematical model incorporates the essential DLT system parameters and calculates energy usage. This approach has the benefit of not needing to run an experimental study, where it would be necessary to run a miner node and measure actual energy consumption. UCL’s model estimates the energy consumption per transaction based on the projected number of validator nodes, their energy consumption, and the network throughput. Their research found a real difference between PoS and PoW. Further, despite the considerably smaller PoS energy consumption, within itself, there are a lot of variances, which could be a problem when those chains escalate in size (Platt et al., 2021).

The carbon footprint generated by mining activities is derived from the energy mix present in the mining node’s locations. Cambridge Center for Alternative Finance has estimated the miners’ locations by working with prominent mining pools across the globe. However, the data is not updated, and the information is not complete (Carter, 2021). It has been estimated that 36% of the energy mix from Bitcoin mining is of sustainable sources (Cambridge, 2021). The actual number is expected to be smaller (Carter, 2021).

Focusing on the energy consumption of mining nodes, the relationship between itself and the carbon footprint of Bitcoin has been recently investigated. The research focus was on the presence of long-term memory on energy consumption, auto-correlation, and mean aversion between the inspected variables by applying different econometric models. It was found that Bitcoin's energy consumption does present a long memory in most cases. The long memory in energy consumption implies that temporary policies will not prevent a lasting effect on consumption. Thus, the paper recommends the creation of permanent legislation to assist in the reduction of carbon emissions in DLTs (Ghosh & Bouri, 2022).

Private entities are aware of the high energy demand from mining Bitcoin, which generates a large carbon footprint. In response, projects such as Bitcoin Net Zero are emerging. The Bitcoin Net Zero project has three main components, the being: partnering with energy providers to source mining facilities, developing energy-efficient hardware, and purchasing carbon offsets to compensate for emissions. The project was made by NTDIC, a New York-based Bitcoin investment firm. The plan also recommends regulations to support this transition. The recommendations can be summarised into 5 different legislative scopes as requiring miners to disclose their energy consumption and emissions, providing incentives for miners to use renewable energy, setting standards for energy efficiency for bitcoin mining hardware, regulating the sale of carbon offsets, and investing in research and development of more sustainable bitcoin mining technologies (NYDIG, 2021).

Apart from considering the energy consumption of the computers, external costs must also be taken into account. The main external sources found in the literature are hardware and cooling costs. Those costs are extreme in PoW systems due to their competitive nature (Klaaßen et al., 2020).

Given that the competitive nature has a significant influence on CO2 emissions, it is pertinent to model it. A way to measure the competitiveness level of a blockchain is to track its hashing rate (Li et al., 2019). In other words, the overall processing power utilized by cryptocurrency to process transactions in a blockchain can be measured by the hashing rate. It can also serve as an indicator of how quickly a Bitcoin miner's apparatus performs the calculations. New hardware allows a miner to calculate hashes faster due to technological enhancements. Consequently, miners are motivated to upgrade their hardware constantly; otherwise, the odds of successfully mining a block will considerably decrease. Further, the study estimates that the life cycle of Bitcoin mining devices is up to 1.29 years. This is because of the high computational demands of mining a new block.

The Digiconomist index also tracks the e-waste generated by the Bitcoin blockchain. Its values are annualized and measured in Kiloton. The growing e-waste generated by the mining cycles is an emerging threat to the environment due to the toxic chemicals and metals that come with unregulated disposal and improper recycling. Consequently, the soil, air and water are polluted (de Vries & Stoll, 2021). The growing e-waste appears to be connected to its price. One indication comes from when Bitcoin e-waste grew beyond 64.4 metric kilotons in early 2021, which coincided with Bitcoin reaching new all-time high prices (de Vries & Stoll, 2021).

Because individuals often are able to sell native tokens in diverse marketplaces, there might be a correlation between the carbon footprint and the price of a DLT system. Studies encountered qualitative reasoning for such a connection (CCRI, 2022). Because of higher prices, individuals searching for profitable ventures build new mining nodes (CCRI, 2022). It is expected that the prices of native tokens affect the carbon footprint of DLTs.

Further research was also made on the main drivers of blockchain prices and its characteristics. The research uses an Auto-Regressive Distributed Lag model to inspect the short and long-term relations. The biggest five public blockchains were considered. It was found that the market beta, trading volume, and volatility have a significant impact on prices. It was also found that the S&P500 index has a weak positive impact in the long run, but in the short run, it becomes negative and insignificant, except for Bitcoin, which has a statistically significant estimate of -0.20 at a 10% significance level. It is also found that the prices between all the blockchains cointegrate (Platt et al., 2021).

This implies that native tokens with a small market price will generate fewer carbon emissions from their operations. Because they have a smaller volume and volatility, on average. However, because they are expected to be cointegrated with other DLTs, in the long term they are expected to have a stable price spread. A type of hybrid blockchain is stable coins, where an institution controls the protocol while its network and application layers are open to the public. In such hybrid DLT, the market price of the native token is fixed to another asset. This is commonly done by pegging it to Euro or Dollar. This will decrease sharply the volatility of the currency and stabilize its market beta. Stable coins are also expected to lessen the competition between crypto miners, which decreases carbon emissions (de Vries et al., 2022). The relationship between price and carbon emission also indicates that DLTs with a low market value have a smaller user base. Subsequently, fewer individuals are maintaining the ecosystem, which indicates a less competitive environment. Unfortunately, there is no available data for hybrid blockchains due to their private nature. A more detailed explanation of public, private, and hybrid DLTs will be provided later in this chapter.

The price variable was also examined by Sarkodie et al. (2022), which investigated how the financial factors of Bitcoin can affect its carbon emissions and vice-versa. This analysis is made by comparing a Vector Auto-Regressive (VAR) model and a dynamic Auto Regressive Distributed Lag (ARDL). In the comparison, they assessed Granger causality cumulative impulse-response relationships and steady-state effects. It was found that their data did not present a serial correlation. Further, no structural brakes were found in their expression residuals. Their ARDL models found that trade volume has a stimulating effect on Bitcoin energy consumption in the long run. However, in the short run, trade volume has a mitigating effect on energy consumption. These results indicate that while trade volume initially reduces energy consumption, it eventually leads to an increase in the long term.

Focusing on the Granger causality results, a unidirectional causality was found from carbon footprint and energy consumption to Bitcoin market capitalization, indicating a conservation interaction. This suggests that a reduction in Bitcoin's energy consumption and carbon footprint could lead to a decline in its market capitalization. On the other hand, bidirectional Granger causality is observed between market price and carbon footprint, as well as trade volume and energy consumption. These findings support the existence of feedback interactions, implying that market price and trade volume can impact Bitcoin's energy consumption and carbon footprint and vice versa. Overall, their results highlight the potential influence of market factors on Bitcoin's energy consumption and environmental impact (Sarkodie et al., 2022).

There is a noticeable ambiguous relationship between innovation and competition regarding DLT's carbon footprint. This is the case, especially in Europe, because of the EU emission trading system. The trading system follows a cap-and-trade approach where corporations can trade limited CO_2 credits in the market. The maximum amount of CO_2 credits decreases over time (European Commission, 2023). With the increasing scarcity of credits, their price is expected to increase. This forces corporations to track their CO_2 emissions and motivates them to decrease their carbon footprint over time. Mining companies, which are known to emit large amounts of CO_2 with their operations, must pass the increasing costs of their emissions to consumers. The way that miners can pass their costs to consumers is by selling native tokens in the marketplace. Consumers who invest in native tokens indirectly help mining companies to buy more CO_2 credits. If miners cannot pass the extra cost to consumers, there will be a loss in competitiveness. Mining companies will have to decrease their operations to meet their carbon quota. Thus, the ability to roll over the costs is directly connected to the native token prices.

While miners can take advantage of native tokens, they also need to take advantage of the rapid innovation in the hardware sector to upgrade equipment. But this comes with an economic and environmental cost: the price of the new equipment and the e-waste generated by the disposal of the old equipment. And there are few ways developed yet to recycle e-waste efficiently. Thus, miners must find a balance between passing the higher costs to consumers -expected to cause losses in miners' market shares- and investing in innovation to maintain their competitive price. The trade-off between competition and innovation becomes eminent because of the EU emission trade system. I anticipate that this trade-off is incorporated into the native token prices and the e-waste generation index.

Costs, competition, and innovation are the main factors driving carbon footprint. Those concepts can be broken down into energy consumption, hashing power, e-waste generation, and prices of native tokens. While those factors are expected to influence costs, competition, and innovation, they can be used to understand the carbon emissions of DLTs. So, I test this expectation using the comparison of data from different DLTs and their emissions. The second hypothesis reads as follows:

"*Energy consumption, hashing rate, e-waste generation, and prices of native tokens have a significant explanatory effect on the blockchain's ecosystem carbon footprint.*"

For this research, the ecosystem will be defined only by the Ethereum and Bitcoin blockchains. Those two blockchains are the most prominent in the public space, with a joint market value of EUR 596bln. Further, according to estimates, Bitcoin uses between 60% and 77% of the world's cryptocurrency assets' power, while Ethereum uses between 20% and 39% (The White House, 2022). However, it would be interesting to consider smaller and private DLTs. This would allow for better segmentation and potential reasoning for carbon footprint drivers in the ecosystem. Especially when there are emerging blockchains with a sustainable focus.

New research on blockchains solutions to the carbon credit market has been conducted (Sipthorpe et al., 2022). In this research, 39 organisations developing blockchain solutions were inspected. They also faced difficulties gathering data because most organisations were unwilling to share their private information. The authors find that the current ecosystem is diverse, fragmented, and relatively immature, with most solutions being early-stage proofs of concept. They identify bottlenecks at various technology readiness levels that indicate critical issues such as scalability, systems integration, and regulation that need to be addressed. They conclude that the growth of blockchain in the carbon markets is significant, but it is often misunderstood and subject to speculation.

A nominal example is the Regen blockchain. Regen is built using a PoS consensus approach and aims to create an ecosystem where individuals can develop impactful ecological projects. Investors can also directly buy carbon credits from farmers. Other verified users can also propose projects for others to support in exchanging carbon credits and other tokens in the ecosystem. Essentially, a public ecological accounting system is created by Regen Ledger. Further, this DLT is designed for the verification of environmental impact claims, agreements, and data. Regen DLT allows registries and credit standards to interact and conduct business with one another.

Companies like Tesla follow a similar approach as Regen. Tesla sells their excess carbon credits to third parties, reaching USD 1.78bln in 2022 (Carbon Credits, 2022). In 2021, Tesla generated more revenue from selling its excess carbon credits than from their electric cars (Carbon Credits, 2023). DLTs like Regen have the potential to bring together enterprises like Tesla with smaller entities and individuals with the benefits of a DLT, which are stated in the OBSF principles.

But to be able to harvest data from a blockchain of this type to derive its energy consumption, one would have to create and maintain an active staking node. To do so, the Ethereum blockchain requires a deposit of 32 ETH, which is its native token (Ethereum, 2023). This amount is equivalent to EUR 47.359,33 on 29th January 2023, according to Binance. For this thesis, it is not feasible to maintain an active staking node. Therefore, I only rely on third-party data.

Blockchains can also be categorised as permissionless, permissioned (private), or both. The permissionless blockchain generally behaves as the one described in Nakamoto's paper, where anyone can join the ecosystem by becoming a node. The node has no restricted rights; it is fully decentralised and allows anyone to have pseudo-anonymity (Nakamoto, 2008). Nowadays, the main purpose of a public blockchain is mining or exchanging currency tokens, such as Bitcoin and Ethereum.

Now, permissioned blockchains do not allow for any type of anonymity. Permissioned blockchains are only partially decentralised. Therefore nodes might have restricted rights. Usually, those DLTs represent ventures undertaken by big corporations or countries aiming to optimise their operation while reducing costs (Amarasinghe et al., 2020). There is also the concept of fully private and consortium blockchains, where the main difference is that the former is governed by only one entity while the latter is managed by a group of organisations. The projects executed within this blockchain segment reach a large spectrum of functionalities: supply chain management, certification of products, bookkeeping, transfer of monetary value, and more (Deloitte 2018; IBM, 2020).

A single organisation or a group controls fully private blockchains. In this case, only users authorised by the central authority can access the blockchain or run a node. Thus, making it a partially decentralised blockchain. Further, the rights between the nodes can vary depending on their clearance level.

Permissionless blockchains are frequently more secure than permission blockchains since it would be difficult for malicious actors to collaborate on the network because there are many nodes to confirm transactions. However, permissionless blockchains have a longer latency when validating new information.

Permissioned blockchains, despite their restrictions, should also have the same carbon emission drivers because, in essence, they are also DLTs (Loreen et al., 2021). Investigating how such restrictions can affect their carbon footprint would be relevant. One restriction could be that all miners must use certain equipment at a specific output power, thus optimising its efficiency considering the carbon footprint. However, no data for private blockchains are available to the public. This research only deals with permissionless blockchains for reasons relating to data accessibility.

2.3.2 Brakes in trends

All types of DLTs have updates which are called forks. Forks can be considered soft or hard. In the majority of DLTs, forks can be proposed by any active node. For the fork to be accepted, the community must accept it. The community usually votes, and the type of vote changes according to the DLT (Alvi et al., 2022).

A hard fork is when the protocol layer has been changed considerably. Such a change consequently creates a new blockchain by requiring all users to upgrade their system to the latest version with no backward compatibility (Frankenfield, 2022). As a result, once a hard fork is executed, a branch of the blockchain is made. One branch follows the previous protocol and a new one follows the updated protocol. An example of a hard fork can be seen on July 20, 2016, when Ethereum performed a protocol update to unmake the effects of a hack and prevent it from happening again (Ethereum, 2023).

Soft forks are also updates in the protocol but allow for backward compatibility. Thus, users are not required to update their system immediately (Frankenfield, 2023). Ethereum and other DLTs created soft forks to implement new functionalities or optimise existing ones. A significant soft fork was made by Ethereum on September 15, 2022, when it changed its consensus approach from PoW to PoS. After the update, Ethereum's carbon footprint shrank, it dropped by more than 95% (Cooling, 2021). Ethereum's forks are listed in appendix A Table 1A.

Updates in the protocol have the potential to significantly change the performance of DLTs (Yang et al., 2020). Moreover, regulation has the potential to change the performance of a DLT system. For example, countries are banning the usage of cryptocurrencies. If a region has a significant user base, removing them from the ecosystem can affect its dynamics. This is the case for China, which before its ban its participation accounted for around 70% of the blockchain ecosystems (Sergeenkov, 2022).

It is pertinent to check if changes in protocol or in the legislative environment are significant to the point of creating structural breaks in the carbon footprint of DLTs. To perform such check the Chow-break test can be employed. Thus, my third hypothesis reads as follows.

"*There is at least one structural break in a blockchain carbon footprint per transaction.*"

If a structural beak is found, the previous hypotheses will be re-accessed. For this research, I will test for one structural break in each blockchain. One break will refer to a soft fork and another to a legislative change. The soft break is the change of consensus approach in the Ethereum blockchain. The legislative change refers to the date China banned cryptocurrencies in their territory. I choose those movements for testing, because they are the ones with the most relevance within the timeframe of my dataset.

In sum, the literature indicates that energy is the main driver for carbon emission in DLTs. New research is starting to take a broader definition of what causes CO2 emission in blockchains, considering hardware circles, market prices, e-waste generation, and hashing power. Figure 2 displays a schematization of the variables that affect the carbon footprint.

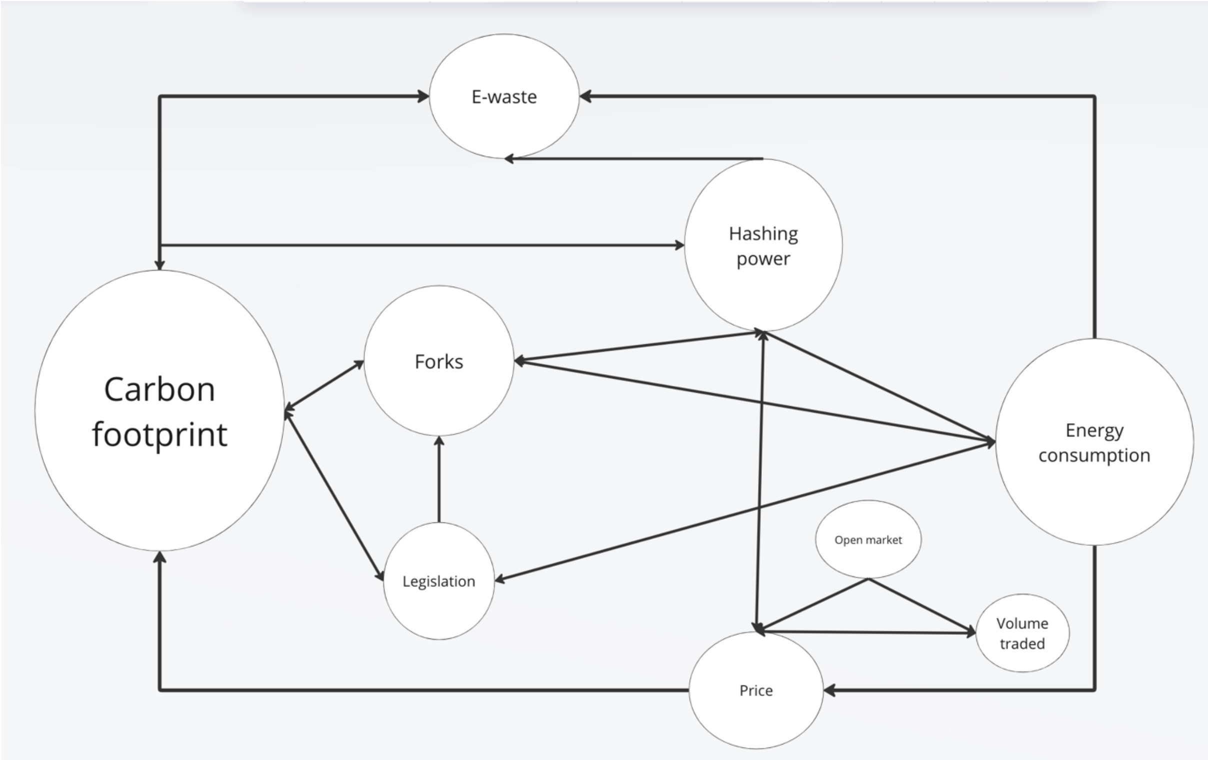

Figure 2 Schematization of variables and interactions that affect the carbon footprint

My goal is to investigate how some of those variables can explain existing measures of carbon emissions from the DLT ecosystem. I would like to also test more on the nuances between private and public blockchains, but there is little to no data available.

CHAPTER 3 Data

In this chapter, I will expand on the sources for the data used in this research, give general statistics for all the variables, and explain the methods applied to ensure that the data has statistical significance.

As mentioned in the Literature, only Bitcoin and Ethereum blockchains will be inspected. Those 2 blockchains incorporate most transactions in the ecosystem and have the highest adoption rate (The White House, 2022). Thus, those two have the potential to represent the public DTL ecosystem.

The gathered data categories are hashing rate, energy consumption, transaction, e-waste, and price. For the Bitcoin blockchain, the hashing rate and the number of transactions were gathered from the Nasdaq database. Nasdaq stores financial information from Ethereum, Bitcoin, and other financial instruments not listed in their exchange. (https://data.nasdaq.com/data/BCHAIN-blockchain).

The energy consumption and carbon emissions were retrieved from the Cambridge Bitcoin Electricity Consumption Index. The market price and e-waste generation were collected from Yahoo Finance and Digiconomics, respectively. The data on Ethereum have overlapping sources, such as Yahoo Finance and Digiconomist. The hashing rate and the number of transactions were sourced via Etherscan.

Further, the carbon emissions were retrieved by CCRI and Kyle McDonald database, and energy consumption by Digiconomist. The carbon emission data from Kyle McDonald goes until the soft fork from Ethereum (2022). After that moment, only CCRI tracks the carbon footprint of the Ethereum blockchain.

To investigate the carbon footprint per transaction in blockchains, I focused on the energy consumption per transaction, hashing power, annualized e-waste generation, their consensus approach, price, and a constant. All the data is presented on a daily basis, the hashing power is presented in Terahashes, carbon emissions in Megatons, e-waste in Kilotons, prices in USD, and energy in Terawatts.

Because Ethereum changed its protocol, a categorical variable was created to consider that change. The e-waste is, in principle, a by-product of maintaining Bitcoins operations. Because Ethereum also follows the same consensus mechanism most of the time, it's reasonable to assume that a similar trend could be found. The carbon footprint and energy per transaction were calculated by dividing the total daily values by the daily transactions. The ecosystem carbon footprint is the summation of Ethereum and Bitcoin's carbon footprint per transaction. The descriptive statistics of the raw data are displayed in Table 1.

Table 1 Descriptive statistics for raw variables

Variables	Mean	Median	S.D.	Max	Min	Skewness	Kurtosis	Obs
CO2 BTC	142.009	143.323	77.443	346.489	0.239	0.067	2.089	1438
Hashing power BTC	7.39×10^{15}	1.67×10^{15}	8.69×10^{15}	5.85×10^{16}	4.71×10^{17}	0.802	2.664	1438
Price BTC	2.37×10^{10}	1.78×10^{10}	1.76×10^{10}	6.76×10^{10}	1.78×10^{4}	0.697	2.130	1438
E-waste	25.755	23.990	7.756	44.230	12.190	0.442	2.648	1438
Energy BTC	0.001	0.001	0.000	0.001	0.000	0.786	2.328	1438
Hashing power ETH	4.26×10^{6}	2.52×10^{6}	3.37×10^{6}	1.13×10^{7}	0.000	0.709	1.999	1438
CO2 ETH	0.009	0.008	0.005	0.021	0.000	0.411	2.622	1438
Price ETH	1.32×10^{9}	6.71×10^{8}	1.29×10^{9}	4.81×10^{9}	1.05×10^{8}	0.838	2.482	1438
Energy ETH	0.000	0.000	0.000	0.000	0.000	1.107	2.764	1438
Consensus ETH	1.063	1.000	0.244	2.000	1.000	3.587	13.869	1438
Ecosystem CO2	142.019	143.332	77.445	346.504	0.252	0.067	2.088	1438

Note. BTC stands for Bitcoin, ETH refers to Ethereum. S.D. means standard deviation and CO2 represents carbon emission. Energy relates to energy consumption. All the data is presented on a daily basis, the hashing power is presented in Terahashes, carbon emissions in Megatons, e-waste in Kilotons, prices in USD, and energy in Terawatts. Consensus ETH is a dummy variable with a value of 1 when ETH follows a PoS approach.

In Table 1, all variables have a similar mean and median, but the hashing rate of Ethereum, which is expected from a normally distributed sample. Considering the SD, 63% of the variables present a value larger than one, which indicates high variance in the sample. The normal skewness is of value zero with a 0.5 difference in both directions. Thus, only three variables are not considered to be skewed, them being e-waste, the carbon footprint of Ethereum and the ecosystem. Further, all variables are positively skewed. Additionally, all variables but the dummy variable for the change in the consensus approach from Ethereum have a kurtosis smaller than 3, thus presenting a thin-tailed distribution.

To further enhance the data quality, outliers from the generated carbon emission in the Bitcoin blockchain were removed. Such analysis was made by plotting the values over time and looking for clear errors in the data. Please find the plotted values from before and after treatment in Figure 1B and Figure 2B appendix B. The summary statistics for all treated variables can be found in Table 2.

Table 2 Descriptive statistics of all cleaned variables

Variables	Mean	Median	S.D.	Max	Min	Skewness	Kurtosis	Obs
CO2 BTC	156.621	154.096	67.838	346.489	50.541	0.220	2.020	1290
Hashing power BTC	7.26×10^{15}	1.67×10^{15}	8.58×10^{15}	5.85×10^{16}	4.71×10^{7}	0.830	2.810	1290
Price BTC	2.36×10^{10}	1.76×10^{10}	1.76×10^{10}	6.70×10^{10}	1.78×10^{4}	0.710	2.150	1290
E-waste	25.736	23.960	7.718	44.230	12.190	0.460	2.680	1290
Energy BTC	0.000	0.000	0.000	0.001	0.000	0.800	2.360	1290
Hashing power ETH	4.21×10^{6}	2.51×10^{6}	3.35×10^{6}	1.13×10^{7}	0.000	0.740	2.060	1290
CO2 ETH	0.009	0.008	0.005	0.021	0.000	0.430	2.680	1290
Price ETH	1.31×10^{9}	6.36×10^{8}	1.28×10^{9}	4.74×10^{9}	1.05×10^{8}	0.840	2.480	1290
Energy ETH	0.000	0.000	0.000	0.000	0.000	1.140	2.850	1290
ETH Consensus	1.064	1.000	0.245	2.000	1.000	3.550	13.610	1290
Ecosystem CO2	156.630	154.104	67.840	346.504	50.548	0.220	2.020	1290

Note. BTC stands for Bitcoin, ETH refers to Ethereum. S.D. means standard deviation and CO2 represents carbon emission. Energy relates to energy consumption. All the data is presented on a daily basis, the hashing power is presented in Terahashes, carbon emissions in Megatons, e-waste in Kilotons, prices in USD, and energy in Terawatts. ETH consensus is a dummy variable with a value 1 when ETH follows a PoS approach.

What leaps out of the page is the big difference between the mean and medium for the price and hashing power for both DLTs. The price from Ethereum also presented the same phenomena. This implies a clustering over time (e.g., heteroskedasticity). For the SD in Bitcoin's energy consumption, Ethereum hashing power and energy consumption was close to 0. The skewness was once again all positive and with the same variables following a normal distribution. The pattern for the kurtosis also follows the trend from the raw data. The largest difference between the raw and cleaned data is the difference between the means and medians.

I checked for the stationarity of all variables by applying unit root tests from Dickey-Fuller (DF) test (1979), Phillips-Perron (PP) test (1988) and the Kwiatkowski-Phillips-Schmidt-Shin test (KPSS) (1992). According to the DF and PP tests, the stationary variables are Bitcoin's hashing power, Bitcoin's energy consumption and Bitcoin's carbon footprint. However, the KPSS test indicated that all variables are non-stationary. An overview of results from the various unit root tests can be found in appendix A, Table 2A, 4A and 6A.

With the results from the KPSS test, I calculated the log differences for all non-stationary variables. I retested the log difference to ensure stationarity. There was insufficient evidence from all the variables' log differences to reject the null hypothesis of stationarity. These results are shown in Table 3A, 5A and 7A.

With the goal of discovering which lags are relevant, I analyzed the autocorrelation (AC) and partial autocorrelation (PAC) for the carbon emission of both blockchains. The results from the AC and PAC calculations can be found in appendix B, Figures 3B - 6B.

Ultimately, all variables were stationary after performing the aforementioned methods. At least one relevant lag was found in the carbon footprints of both blockchains. The descriptive statistics of the treated variables can be found in Table 3 below.

Table 3 Descriptive statistics of treated variables

Variables	Mean	Median	S.D.	Max	Min	Skewness	Kurtosis	Obs.
CO2 BTC	-0.001	-0.011	0.123	0.585	-0.472	0.390	3.715	1289.000
Hashing power BTC	-0.001	0.000	3.550	16.183	-16.152	0.107	5.508	1289.000
Price BTC	0.010	-0.001	0.387	13.814	-0.179	35.302	1260.331	1289.000
Energy BTC	-0.001	-0.001	0.019	0.120	-0.078	0.587	6.665	1289.000
CO2 ETH	-0.001	-0.009	0.124	0.559	-0.564	0.318	3.770	1289.000
Hashing power ETH	-0.000	-0.000	0.044	1.405	-0.064	25.022	797.025	1290.000
Price ETH	0.018	-0.003	0.639	22.683	-0.580	34.690	1230.862	1289.000
Energy ETH	-0.002	-0.002	0.051	0.380	-0.249	0.703	8.823	1289.000
E-waste	0.006	-0.004	0.202	5.680	-0.562	20.261	533.041	1289.000

Note. BTC stands for Bitcoin, ETH refers to Ethereum. S.D. means standard deviation and CO2 represents carbon emission. Energy relates to energy consumption. All the values are calculated with log differences.

In contrast to before, the treated variables do not present a considerable difference between the mean and median. Another disparity is the large kurtosis and skewness in the Bitcoin's price log difference. This is somewhat expected for the Bitcoin market price is known to have such properties (Vieira & Laurini, 2022). Ethereum characteristics are similar to before treating the data.

Another relevant statistical measure that is relevant for answering the hypotheses is the correlation between the variables. For it gives insight into the priority of the exogenous variables against carbon emission of blockchains. Bitcoin energy has the highest correlation followed byBitcoin native token prices, e-waste and Bitcoins hashing levels. The correlation summary for Bitcoin and Ethereum is displayed in Tables 4 and 5. This goes in line with my expectations because energy carbon emission can be considered a by-product of the energy consumption of a process. For Bitcoin's price, the high correlation might be a result from the increased incentive for people to start mining or upgrading their equipment. A similar rationale can be derived for the e-waste variable.

Table 4 Bitcoins correlation matrix

Variables	CO2 BTC	Hashing power BTC	Price BTC	E-waste	Energy BTC
CO2 BTC	1	0.5462	0.7940	0.6355	0.8079
Hashing power BTC	0.5462	1	0.5051	0.5544	0.5479
Price BTC	0.7940	0.5051	1	0.3926	0.6627
E-waste	0.6355	0.5544	0.3926	1	0.6216
Energy BTC	0.8079	0.5479	0.6627	0.6216	1

Note. BTC stands for Bitcoin, and CO2 represents carbon emission. Energy refers to energy consumption. The variables used are in absolute terms.

This correlation matrix provides valuable insights into the relationships among key variables associated with Bitcoin. Hashing power exhibits a moderate positive correlation with carbon emissions, indicating that increased computational power in Bitcoin mining is linked to higher carbon emissions. Similarly, carbon emissions show a strong positive correlation with energy consumption, suggesting that greater energy usage in Bitcoin mining leads to increased carbon emissions.

Bitcoin's price demonstrates a strong positive correlation with both carbon emissions and energy consumption. This implies that as the price of Bitcoin rises, there is a corresponding increase in carbon emissions and energy consumption within the Bitcoin network. Additionally, e-waste displays moderate positive correlations with carbon emissions, hashing power, and energy consumption, highlighting the environmental implications of electronic waste resulting from Bitcoin mining activities. Further, the variables with the lowest correlation with Ethereum carbon emission is the e-waste. The correlation matrix for Ethereum can be found in Table 5.

Table 5 Ethereum correlation matrix

Variables	CO2 ETH	Hashing power ETH	Price ETH	E-waste	Energy ETH
CO2 ETH	1	0.9318	0.6104	0.2064	0.9019
Hashing power ETH	0.9318	1	0.7822	0.4579	0.9428
Price ETH	0.6104	0.7822	1	0.4619	0.6716
E-waste	0.2064	0.4579	0.4619	1	0.4447
Energy ETH	0.9019	0.9428	0.6716	0.4447	1

Note. ETH stands for Ethereum, and CO2 represents carbon emission. Energy refers to energy consumption. The variables used are in absolute terms.

Table 6 presents a correlation matrix between Bitcoin and Ethereum regarding CO2 emissions, hashing power, price, and energy consumption. The findings indicate a moderate positive correlation between CO2 emissions and hashing power between Bitcoin and Ethereum. There is a strong positive correlation in prices, suggesting a close connection in market values. Additionally, Bitcoin and Ethereum demonstrate a relatively high positive correlation in energy consumption. These findings highlight the interdependency and interconnectedness of Bitcoin and Ethereum in terms of their environmental impact, computational power, market value, and energy usage.

Table 6 Ethereum vs Bitcoin correlation matrix

Variables	CO2 BTC	Hashing power BTC	Price BTC	Energy BTC
E-waste	0.243	0.443	0.440	0.445
CO2 ETH	0.551	0.699	0.774	0.605
Hashing power ETH	0.321	0.492	0.529	0.412
Price ETH	0.501	0.689	0.922	0.521
Energy ETH	0.710	0.831	0.831	0.827

Note. ETH stands for Ethereum, CO2 represents carbon emission and BTC for Bitcoin. Energy refers to energy consumption. The variables used are in absolute terms.

In sum, the majority of the data follows a normal distribution, considering its statistical values. Also, all the treated variables passed the unit root tests, allowing for statistical analysis (Dickey & Fuller, 1979). Lastly, there is at least one highly correlated variable against the CO2 from its respective blockchain.

CHAPTER 4 Methods

I will describe the methodology for each hypothesis in detail, together with robustness check suggestions. Those checks are for the case some expectations are proven to be true.

4.1 Hypothesis 1

The first hypothesis reads:

"There is a statistical difference between the average carbon emission per transaction across different DLTs."

Once all the data treatment is made, as described in the data section, the first hypothesis can be tested. To do so, a Two-Sample-T-test is employed between the log difference of the carbon footprint of Ethereum and Bitcoin. The Two-Sample-T-test is defined in the following expression.

[1] Two-Sample-T-test

$$t - value = \frac{(x_1 - x_1) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

A two-sample t-test is commonly used to evaluate whether the difference between the means of two groups is statistically significant or due to chance. The goal is to discover if there is a statistical difference between the samples. I expect large differences in the carbon emissions per transaction within both blockchains. Especially because after a certain period, Ethereum changed its consensus approach, as explained in the background literature.

4.2 Hypothesis 2

To construct a model that has as a help explain the carbon footprint for the blockchain ecosystem per transaction, a bottom-up approach is used. The general explanatory model is defined in expression [2] The observed carbon footprint of the ecosystem is the summation of the CO2 emissions from Bitcoin and Ethereum, as mentioned in the Data section. The second hypothesis is as follows:

"*Energy consumption, hashing rate, e-waste generation, and prices of native tokens have a significant explanatory effect on the blockchain's ecosystem carbon footprint.*"

[2] General ecosystem model

$$Carbon\ footprint\ ecoystem_t = \alpha_t + \beta_t Carbon\ footprint\ of\ Ethereum_t + \beta_t Carbon\ footprint\ Bitcoin_t + \varepsilon_t$$

Where the carbon footprint of Ethereum and Bitcoin are independent models that are derived following expression [3], α is a constant and the ε represents the error term. The derivation process is explained below.

For the individual carbon footprint function, a bottom-up approach is considered. Thus, I start with a simple model by fitting a linear regression and then evaluating its fit by checking the coefficient significance. Further, I access heteroscedasticity and autocorrelation in the model. Those tests are made by applying the Breusch-Pagan-Cook-Weisberg and the Lagrange Multiplier test, respectively. If those characteristics are found in the model, Newley West (NW) standard errors will be used. After checking the fitting of the model, the polynomials of the same variable will be added up to the fourth dimension. Human behaviour biases such as risk aversion and the non-linear behaviour of renewable energy generation are the main reasons to assume a non-linear relation between the inspected variable and the exogenous one. The order in which the variables will be added to the model will have both quantitative reasoning, by inspecting their correlation, and qualitative, where there must be a relevant rationale for the explanatory variable to influence the carbon footprint of the blockchains. The bottom-up approach order is as follows: Energy consumption of the blockchain, hashing power, e-waste generation, type of consensus mechanism – when applicable -, and the native token price. The general model for each blockchain can be defined as an expression [3].

[3] General model

$$Carbon\ Emission_t = \alpha + \sum_{s=1}^{4} (\beta_{t,s} X_{t,s}^1 + \beta_{t,s} X_{t,s}^2 + \beta_{t,s} X_{t,s}^3 + \beta_{t,s} X_{t,s}^4) + \varepsilon$$

Where “s” stands for the relevant variables that explain the carbon emission, α is a constant, β is the correlation coefficient, and the ε represents the error term. To further expand on the proposed model and its implications, I calculated the VAR model. VAR models are a class of time-series models used to analyse the dynamic relationship among multiple variables. VAR model considers interdependencies within the variables. The interdependencies refer to the fact that the behaviour of each variable is influenced by the values of other variables in the system. The variables are not considered in isolation but are seen as interconnected and affecting each other. In a VAR model, each variable is modelled as a function of its own lagged values and the lagged values of other variables in the system (Sims, 1980). A VAR model also allows for inspecting Granger causality, which evaluates whether the previous values of one variable may be used to predict the future values of another variable. To estimate and VAR model, all variables must be tested for unit root. As mentioned in the Data section, all variables were inspected for it using the DF test, and the necessary changes were made to satisfy the data stationarity requirements. Thus, the variables under analysis refer to the ones described in Table 3. Further, the optimal number of lags for the expression must be estimated. In this case, I minimise the Bayesian Information Criterion (BIC). The BIC is defined in expression [4] (Schwarz, 1978).

[4] BIC

$$BIC = kln(n) - 2\,ln(L)$$

Where k is the number of parameters in the model, n the number of observations and L the model likelihood function BIC tends to underestimate the optimal amount of lag. (Bruns & Stern, 2018) which can be a problem because the Granger causality test in a VAR model is very sensitive to the number of lags (Zapata & Rambaldi, 1997). However, underestimation of the number of lags provides a good fit to the data while avoiding overfitting. In this case, the frequency of the VAR model is daily observations. The general form of a VAR model can be written as an expression [5] below.

[5] VAR

$$Y_T = \alpha + \sum_{i=1}^{k}(\beta_i X_{11}) + \sum_{j=1}^{k}(\beta_j X_{j2}) + \sum_{m=1}^{k}(\beta_m X_{m3}) + \varepsilon_t$$

Where Y_T is a vector of endogenous variables at time t, α is a constant vector, and k is the number of lags. The β is the short-run dynamic coefficient, and the ε represents the residual in the equation. After deriving the VAR an Impulse Response Function (IRF) will be performed. The IRF examines how variables respond to a one-time shock. It quantifies the magnitude, duration, and direction of the response over time. The impulse response function shows the accumulated impact of the shock and helps understand system dynamics and effects.

After observing the IRF, I will test if the explanatory variables granger causes the dependent one. The Granger causality test is a statistical technique for assessing whether a time series may be used to predict another.Being able to explain carbon emissions in different blockchains is a crucial point. It will allow a better understanding of which segments of this emerging technology significantly impact carbon emissions. This information could assist interested parties in minimizing their carbon emission from other DLS.

If endogeneity, and particularly the simultaneity problem, is found in the models, the Two-Stage-Least Square method (2SLS) will be applied. Endogeneity issues are frequently faced in econometrics, where dual causation is encountered. To execute the 2SLS, two steps are followed.

In the first step, the endogenous variable in the model is estimated using an Instrumental Variable (IV). An IV correlates with the endogenous variable but not with the error term in the regression equation. The endogenous variables from the first stage's estimated values are employed as the independent variable in the second stage's regression equation(s) for the dependent variable. The endogenous variables' estimated values are referred to as fitted values. Then, in the regression equation(s) for the dependent variable, the fitted values are substituted for the original endogenous variables.

The 2SLS approach aids in removing this bias. The 2SLS approach offers reliable estimates of the causal relationship(s) between the variables of interest by using an IV to estimate the endogenous variable(s). The 2SLS approach, however, necessitates the discovery of a reliable instrumental variable, which can be challenging. If necessary, I will test all the other variables not used in the model as n IV.

After checking for Granger causality, it is pertinent to assess the cointegration between the variables. This test is made by performing the Johansen tests for cointegration. This statistical test is used to determine the presence and number of long-term relationships among a set of variables. These tests involve estimating a VAR model and examining eigenvalues to assess the stationarity of linear combinations of the variables. A Vector Error Correction Model (VECM) will be employed if cointegration is found. The VECM is an extension of the VAR model that considers not only its short-term dynamics correlations but also long-term equilibrium relationships. This is done by implementing an error correction term (Granger, 1969). The VECM model is in expression [6].

[6] VECM

$$\Delta \mathrm{Yijt} = \beta 0\mathrm{ij} + \sum \beta\mathrm{ijl}\ \mathrm{k}\ \mathrm{l} = 1\ \Delta \mathrm{Yijt} - \mathrm{l} + \lambda\mathrm{ijECTjt} - 1 + \mathrm{uijt}$$

Where $\Delta Yijt$ represents the change in the ith endogenous variable for the jth subject at time t, $\beta 0ijt$ is the constant term, $\sum \beta ijl\ k\ l$=1 represents the lagged coefficients up to the kth lag, λij represents the long-run equilibrium relationship between the ith endogenous variable for the jth subject and the error correction term $ECTt-1$, and $uijt$ represents the error term or residual. $ECTt-1$ is the lagged residual of the cointegrating equation.

The main difference between VECM and VAR models lies in their treatment of cointegration. VECM explicitly models long-run equilibrium relationships among variables through cointegration, capturing both short-run dynamics and long-run adjustments. It includes an error correction term to measure the speed of adjustment towards equilibrium. In contrast, VAR assumes stationary variables without cointegration, focusing solely on short-run dynamics. VECM requires estimating cointegrating relationships, error correction terms, and short-run dynamics, while VAR estimates autoregressive relationships without considering cointegration. VECM is suitable when cointegration is present, while VAR is used for analysing short-run dynamics without explicit consideration of long-run equilibrium.

4.3 Hypothesis 3

To inspect the third hypothesis, the Chow-break test was implemented. Hypothesis 3 reads:

"*There is at least one structural break in a blockchain carbon footprint per transaction.*"

The Chow-break test will be regressed using the model that presented the highest significance in hypothesis 2 after the bottom-up approach. Breaks will also be inspected considering only the energy consumption of each blockchain and a constant, as displayed in expression [7]. The Chow-break test for specific periods is defined as in expression [8]**.**

[7] Regression to check for structural breaks.

$$Y_T = \alpha + \beta_i Energy\ Consumption + \varepsilon_t$$

[8] Chow-break test

$$F = \frac{\frac{(RSS_R - RSS_U)}{(k+1)}}{\frac{RSS_U}{DoF}}$$

Where the RSS stands for the sum of squared residuals, the R relates to the restricted model, which only takes into account the period after the break and U to an unrestricted model, thus accounting for the whole sample. Further, the K stands for the number of coefficients in the model plus one, which counts for the constant. Also, DoF means degrees of freedom. The null hypothesis is that there are no structural break at the restriction moment.

The dates selected for inspecting structural breaks are chosen by qualitative reasoning, such as when China banned cryptocurrency, or Ethereum changed its protocol from proof of stake to proof of work. I chose those two test moments because they have a clear reasoning for creating a shock in the carbon emission of the blockchain's ecosystem. China banning crypto mining forced miners to change their location. The ban also changes the energy mix used in the ecosystem because miners moved to other regions with different energy sources. This is the case because China has 30% of its energy produced by renewable sources, which is considerably high compared to other countries in the region, and there where a concentration of Bitcoin miners in the region. At the same time, other countries have worse ratios. Further, I will apply the Quandt-Likelihood Ratio (QLR) to look for other structural breaks. QLR is defined as expression [9], where $F(\tau)$ is the statistic computed over a range of eligible break dates $\tau 0 \leq \tau \leq \tau 1$.

[9] QLR model

$$QLR = max[F(\tau 0), F(\tau 0 + 1), \dots, F(\tau 1)]$$

Performing the QLR analysis is important because it essentially performs the Chow-break test for each possible t. The test being positive alone does not make that moment a significant break. Qualitative reasoning is also necessary to make the break relevant. My qualitative reasoning for the potentially found breaks can be found in the Discussion Section, touching upon what it may mean for future regulation and research. Lastly, I will re-examine the previous hypothesis considering the suggested breaks.

4.4 Checking for predictive power

The primary goal of this section is to explain how I will examine the predictive power of the proposed models. The main approach is to test the Root Mean Squared Error (RMSE) from the proposed models against the benchmark by applying the Diebold-Mariano (DM) test. The benchmark is an Auto-Regressive (AR) model with one lag. A mathematical definition of the benchmark is shown in the expression [10].

[10] Benchmark

$$Ecosystem\ Carbon\ footprint_{\ t} = \alpha_t + \ \beta_{t-1} * \ EcosystemCarbon\ footprint_{\ t-1,} + \ \varepsilon_t$$

Where α is a constant, β is the correlation coefficient, and the ε represents the error term. This approach will be taken for both individual blockchains and for the ecosystem. A plot of three benchmarks over time is displayed in Figure 7B - 9B in appendix B.

The benchmarks are tested against the proposed models from hypothesis 2 three times. To minimise overfitting in the forecasts, pseudo-out-of-sample forecasts will be utilised. To do so, I will employ a rolling window. Also, I will use 30 days to train the model. Thus, the first forecast will be for the 31st day. The proposed model is an autoregressive model with multiple variables whose lags are fixed over time.

To assess if the benchmark is worse than the suggested model, I will perform the DM test. The DM test assesses whether one model significantly outperforms the other. The test is based on the difference in forecast errors between the models. By comparing the RMSE and conducting a hypothesis test, it determines if one model is significantly more accurate than the other. Thus, if the proposed model has a statistically smaller RMSE, the same outperforms the benchmark. The proposed models will be derived by the end of the third hypothesis.

CHAPTER 5 Results

In this chapter, I lay out the results of the tests described in the methodology and comment on their implications. I investigate each hypothesis separately and comment on the results.

5.1 Hypothesis 1

This hypothesis aims to investigate whether diverse protocols have a different average carbon emission per transaction, following the expectations from the literature. As explained in the methodology, a two-sample T-test was employed to explore the possibility. The two-sample t-test for both perspectives can be found in Table 7. Despite their structural differences, no significant difference was found between Bitcoin and Ethereum's average log difference in CO2 emissions per transaction, given a confidence level of 95%. Considering the absolute values, which represent the actual levels of CO2 emissions, the null hypothesis of equal averages is rejected. Further, the test also estimates that Bitcoin's average is larger than that of Ethereum. Those results imply that Bitcoins and Ethereum's carbon footprints are fundamentally different and follow a similar change over time, which is in line with the expectations from Hypothesis 1.

Table 7 Two-sample T-test for the average carbon footprint

Null Hypothesis	P-value
CO2 BTC = CO2 ETH	0.00
CO2 BTC Log Difference = CO2 ETH Log Difference	0.29

Note. BTC stands for Bitcoin, and ETH for Ethereum. CO2 relates to carbon emission. The log difference represents the change in the CO2 levels.

5.2 Hypothesis 2

While the first hypothesis inspected the absolute difference in the average carbon footprints, the second hypothesis aims to explain their carbon emission trends. The variables used for hypothesis 2 are: energy consumption per transaction, hashing power, e-waste generation, type of consensus mechanism -if applicable-and price. As explained in the methodology, I tested the heteroskedasticity and autocorrelation for the relevant variables during the bottom-up approach. The tables from the bottom-up approaches (40 models for BTC and ETH) can be found in appendix A, Tables 8A - 24A. Next, Tables 25A and 26A present the heteroskedasticity in the data sample and the autocorrelation within the first lag. While covering this data in the appendix, the following paragraphs elaborate on models for Ethereum, Bitcoin and the ecosystem.

5.2.1 Hypothesis 2 Ethereum

Starting with the Ethereum blockchain, energy consumption is found to have a significant correlation with a linear and non-linear approach. The final formula with the OLS and NW standard errors for the Ethereum Blockchain, based on Tables 12A – 16A from the appendix, can be found in Table 8. All variables were estimated to be significantly different from 0. The first, second, and third polynomials of energy consumption were found relevant at a 5% significant level. No other variable proved applicable when explaining the changes over time in Ethereum's carbon emissions.

Table 8 Model for Ethereum carbon emissions

Variable	(1) OLS	(2) NW
Energy ETH	1.047***	1.047***
	(0.199)	(0.227)
Energy ETH 2nd Polynomial	-0.549**	-0.549***
	(0.173)	(0.114)
Energy ETH 3rd Polynomial	0.0723*	0.0723***
	(0.0281)	(0.0132)
Constant	0.0223	0.0223
	(0.018)	(0.019)
Observations	1289	1289
R^2	0.025	-

Note. Standard errors are in parentheses. The OLS column stands for the model using the Ordinary Least Square standard errors. The NW column represents the model applying Newey-West standard errors. ETH stands for Ethereum. CO2 represents carbon emission. Energy relates to energy consumption. $^{*}p < 0.05$. $^{**}p < 0.01$. $^{***}p < 0.001$.

Only the first three polynomials of energy consumption from Ethereum are statistically relevant. Thus, according to the model, they incorporate all the information from the other variables. This was expected, given that the energy consumption, in some way, represents the other variables. For instance, information about the price is incorporated into the energy consumption. When it is profitable to mine the native token, more miners activate their nodes in the expectation of making a profit. Also, the higher the polynomial is, the smaller the coefficient is because the second polynomial decreased only by half and became negative. This implies that the second polynomial has a more significant effect than the first for large values and the opposite for smaller values. So, the negative effect is dominant when there is a large log difference.

The model infers that minor daily differences increase, medium differences decrease, and large ones increase the carbon emissions of the Ethereum blockchain, as visualised in

Figure 3. The third polynomial is six times smaller than the second and positive. This allows for a higher likelihood of a positive trend forming for large values, which is necessary given the observed history of carbon footprints from various blockchains. Besides that, the R squared of 0.025 indicates that the model does not have a strong explanatory power.

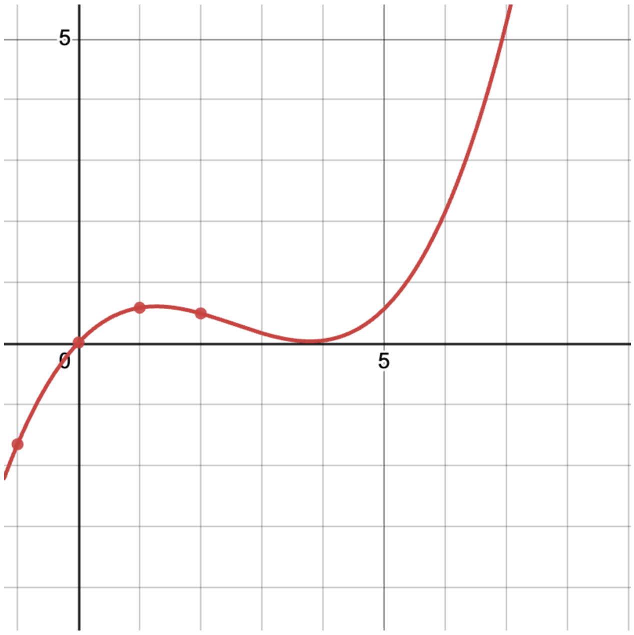

Figure 3 Visualization of model for Ethereum's carbon emissions

To test for causality between the variables of energy consumption and carbon emissions of Ethereum, I performed the Granger causality test, based on the calculated VAR (as explained in the methodology). I considered the optimum number of lags based on the BIC criteria for Ethereum's VAR (see Table 27A in appendix A). From the maximum number of six lags, the optimum number of lags is two. In expression [11], the VAR model for Ethereum is presented.

[11] VAR Estimation Equation Ethereum

Variables	CO2_ETH	ENERGY_ETH1	ENERGY_ETH2	ENERGY_ETH3
R-squared	0.985	0.607	0.889	0.942

$$Y_t = \begin{bmatrix} -0.290 & 0.031 & -0.482 & 0.201 \\ -0.309 & 0.056 & 1.144 & -0.191 \\ -0.655 & 0.552 & 6.437 & -1.090 \\ -3.390 & 2.319 & 36.637 & -6.143 \end{bmatrix} Y_{t-1} + \begin{bmatrix} -0.017 & -0.238 & 0.242 & 0.000 \\ 0.008 & -0.336 & -0.458 & 0.129 \\ 0.002 & -0.263 & -2.671 & 0.559 \\ 0.006 & -1.252 & -16.628 & 3.382 \end{bmatrix} Y_{t-2} + \begin{bmatrix} -0.001 \\ -0.005 \\ -0.014 \\ -0.132 \end{bmatrix}$$

In the short term, the R-squared for all variables is above 0.5, which signals that, more often than not, the regression correctly represents the relationship. The R-squares for the CO2 and energy in its second and third polynomials are close to 1. This indicates a potential overfitting of the data, dual causality, or other biases. Anyhow, the p-values for the equations are statistically significant, meaning that the lagged values of the variables help predict current values. It is also noticeable that the magnitude for the coefficients is not larger than 0.5 in the carbon footprint Ethereum vector.

To better understand how the variables relate, I performed the Impulse-Response Function. What stands out is that the impulses do not generate a permanent change in their trend for all the cases. Also, the carbon emissions of Ethereum do not pose any significant short-term response when receiving the impulse from the other variables. The IRF function is shown in Figure 4.

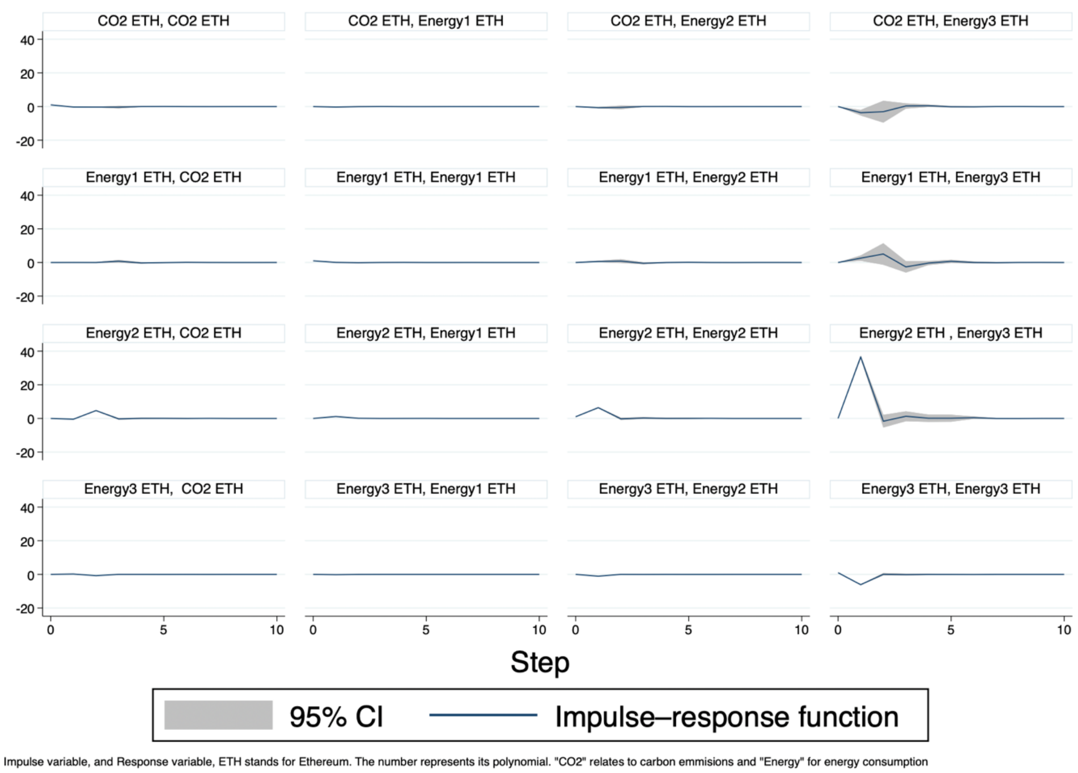

Figure 4 Results from IRF for Ethereum

The coefficients with the p-value for Ethereum's Granger causality test are shown in Table 9. When checking for Granger causality, all regressions have a p-value of 0.000. Thus, there is a two-way Granger causation, implying simultaneity, which can be caused by measurement errors or omitted variable bias (Antonakis et al., 2014).As mentioned in the methodology, there are ways to deal with such problems, such as performing a 2SLS method.

However, it was impossible to discover a reliable IV, which is necessary for the 2SLS method. The tests for IV are shown in Table 28A, appendix A.

Table 9 Granger causality for Ethereum

Y-variable	Exclude	DF	P-value
CO2 ETH	Energy ETH first polynomial	1	0.000
CO2 ETH	Energy ETH second polynomial	1	0.000
CO2 ETH	Energy ETH third polynomial	1	0.000
CO2 ETH	All	3	0.000
Energy ETH first polynomial	CO2 ETH	1	0.000
Energy ETH first polynomial	Energy ETH second polynomial	1	0.000
Energy ETH first polynomial	Energy ETH third polynomial	1	0.000
Energy ETH first polynomial	All	3	0.000
Energy ETH second polynomial	CO2 ETH	1	0.005
Energy ETH second polynomial	Energy ETH first polynomial	1	0.000
Energy ETH second polynomial	Energy ETH third polynomial	1	0.000
Energy ETH second polynomial	All	3	0.000
Energy ETH third polynomial	CO2 ETH	1	0.000
Energy ETH third polynomial	Energy ETH first polynomial	1	0.000
Energy ETH third polynomial	Energy ETH second polynomial	1	0.000
Energy ETH third polynomial	All	3	0.000

Note. ETH stands for Ethereum. CO2 represents carbon emissions. Energy relates to energy consumption. DF means degrees of freedom.

Because a dual causality is present in the data, it is relevant to test for cointegration. The first step is checking for stationarity in the model's residuals. Three stationarity tests, DF, PP and KPSS were performed, the results can be found in Table 29A in appendix A. Only the KPSS test indicated non-stationary in the residuals, giving a reason for this research to continue investigating cointegration within the variables.

With the Johansen-Juselius test for cointegration, I found enough evidence to reject the null hypothesis of no cointegration up to and including the first rank (see Table 30A in appendix A). This further suggests a long-term relationship between the variables. In further depth, through a VECM, I modelled the cointegration behaviour. The VECM results in the estimation displayed in expression [12]. The R-squared from the vectors lay above 0.5, with only one smaller than 0.9. This suggests that the model represents the endogenous variables' relationships in the short and long run.

Further, the cointegration equation with CO2 as endogenous variables found that energy in its first and third polynomials has a negative long-term relationship. In contrast, the energy second polynomial has a positive one. But because the CE vector has a negative value, the coefficients in the cointegration equation can be interpreted with the opposite sign. Therefore, their long-term coefficient arrangement follows a similar trend encountered in the regressions from the VECM equation. The same pattern is not found in their short-term coefficients within the VECM.

[12] VECM Ethereum Full Sample

Variables	CO2_ETH	ENERGY_ETH1	ENERGY_ETH2	ENERGY_ETH3
R-squared	0.989	0.565	0.905	0.957

$$\mathrm{Yt} = \begin{bmatrix} -1.559 \\ -0.886 \\ -1.357 \\ -6.715 \end{bmatrix} \mathrm{CE}_{t-1} + \begin{bmatrix} 0.022 & -0.231 & -0.603 & 0.048 \\ -0.034 & -0.292 & 0.786 & -0.246 \\ -0.109 & -0.381 & 5.305 & -1.139 \\ -0.356 & 1.274 & \mathbf{34.394} & -7.007 \end{bmatrix} \mathrm{LD}_{l} + \begin{bmatrix} 0.000 \\ 0.000 \\ 0.000 \\ 0.000 \end{bmatrix}$$

Variable	CE
CO2 ETH	1***
Energy ETH 1th Polynomial	-0.619*** (-0.167)
Energy ETH 2nd Polynomial	0.802*** (0.41)
Energy ETH 3rd Polynomial	-1.009*** (0.008)
Constant	0.001

5.2.2 Hypothesis 2 Bitcoin

For the Bitcoin carbon footprint model, 22 regressions were considered. Based on Bitcoin's regressions from Tables 8A – 11A in the appendix, I derived the model in Table 10.

Table 10 Model for BTC carbon emission

Variable	(1) OLS	(2) NW
Energy BTC	0.969^{***}	0.969^{***}
	(0.006)	(0.007)
Energy BTC 4th Polynomial	0.461^{**}	0.461^{**}
	(0.163)	(0.164)
Constant	-0.000	-0.000
	(0.001)	(0.001)
Observations	1289	1289
R^2	0.950	

Note. Standard errors are in parentheses. The OLS column stands for the model using the Ordinary Least Square approach. The NW column represents the model applying Newey-West standard error. BTC stands for Bitcoin. CO2 represents carbon emission. Energy relates to energy consumption. $^{*}p < 0.05$. $^{**}p < 0.01$. $^{***}p < 0.001$.

Considering Bitcoin's carbon emissions, energy is highly relevant linearly and in its fourth polynomial. Considering OLS standard errors, the second polynomial proves relevant. The coefficient was positive but close to zero. When accounting for heteroskedasticity and autocorrelation, the second polynomial coefficient became irrelevant. The same happened to the first polynomial of e-waste. Different from hashing power, e-waste presented a large coefficient of 0.079. Other variables did not have significant coefficients.

For Bitcoin, both coefficients positively correlate with the carbon emissions per transaction, which aligns with the literature. Further, because the constant is not significant, the intercept is zero. Implying that no underlying constant trend helps explain the inspected variable. For Bitcoin, the R-squared is extremally high, reaching a level of 95%. An R-squared of this magnitude suggests overfitting of the data or dual causality.

To further understand which added value the controlled variables bring, I calculated VAR after considering their first six lags as recommended by the BIC. The output of the BIC is displayed in appendix A, Table 31A. The VAR tested if the model adds information in the short run while considering the lag of the exogenous variable. The VAR model is displayed in expression [13]. The optimum number of lags is 6.

[13] VAR Estimation Equation Bitcoin

Variables	CO2 BTC	Energy1 BTC	Energy4 BTC
R-squared	0.449	0.442	0.17

$$Y_t=\begin{bmatrix}0.278 & -0.918 & 1.352\\ 0.182 & -0.802 & 1.194\\ 0.001 & -0.009 & 0.370\end{bmatrix}Y_{t-1}+\begin{bmatrix}-0.286 & -0.456 & -1.65\\ -0.348 & -0.378 & -0.285\\ 0.002 & -0.005 & -0.115\end{bmatrix}Y_{t-2}+\begin{bmatrix}-0.085 & -0.565 & 0.494\\ -1.178 & -0.459 & 0.436\\ 0.007 & -0.007 & 0.024\end{bmatrix}Y_{t-3}$$

$$+\begin{bmatrix}-0.064 & -0.553 & 0.359\\ -0.085 & -0.524 & 0.230\\ 0.004 & -0.004 & -0.002\end{bmatrix}Y_{t-4}+\begin{bmatrix}-0.146 & -0.391 & -1.255\\ -0.186 & -0.339 & -0.986\\ 0.007 & -0.007 & -0.011\end{bmatrix}Y_{t-5}+$$

$$\begin{bmatrix}-0.078 & -0.301 & 0.968\\ 0.142 & -0.344 & 1.118\\ 0.003 & -0.003 & -0.005\end{bmatrix}Y_{t-6}+\begin{bmatrix}-0.005\\ -0.005\\ -0.001\end{bmatrix}$$

The VAR's R-squared for CO2 and Energy (1) are a moderate fit and statistically significant. This is because they are close to 0.5. Bitcoin's CO2 vector has a negative relationship with energy in its first polynomial for all lags. If a constant and large positive change in energy consumption occurs, the VAR model suggests an increase in the carbon footprint. If a small positive constant change is found, a negative change in the carbon footprint could be observed. For negative values, the change is positive once again. Therefore, the carbon footprint vector for Bitcoin would present a convex shape.

From the VAR for Bitcoin, the impulse and responses between the variables were calculated. The results from IRF are displayed in Figure 5. From this figure, it is possible to infer that the fourth polynomial of energy consumption in the Bitcoin blockchain has a long-term effect on the carbon emission of this estimation. Now, the opposite does not generate a response impulse. This implies that legislation in the energy consumption of DLTs can have a permanent effect on the carbon footprint of the same DLT but not the opposite.

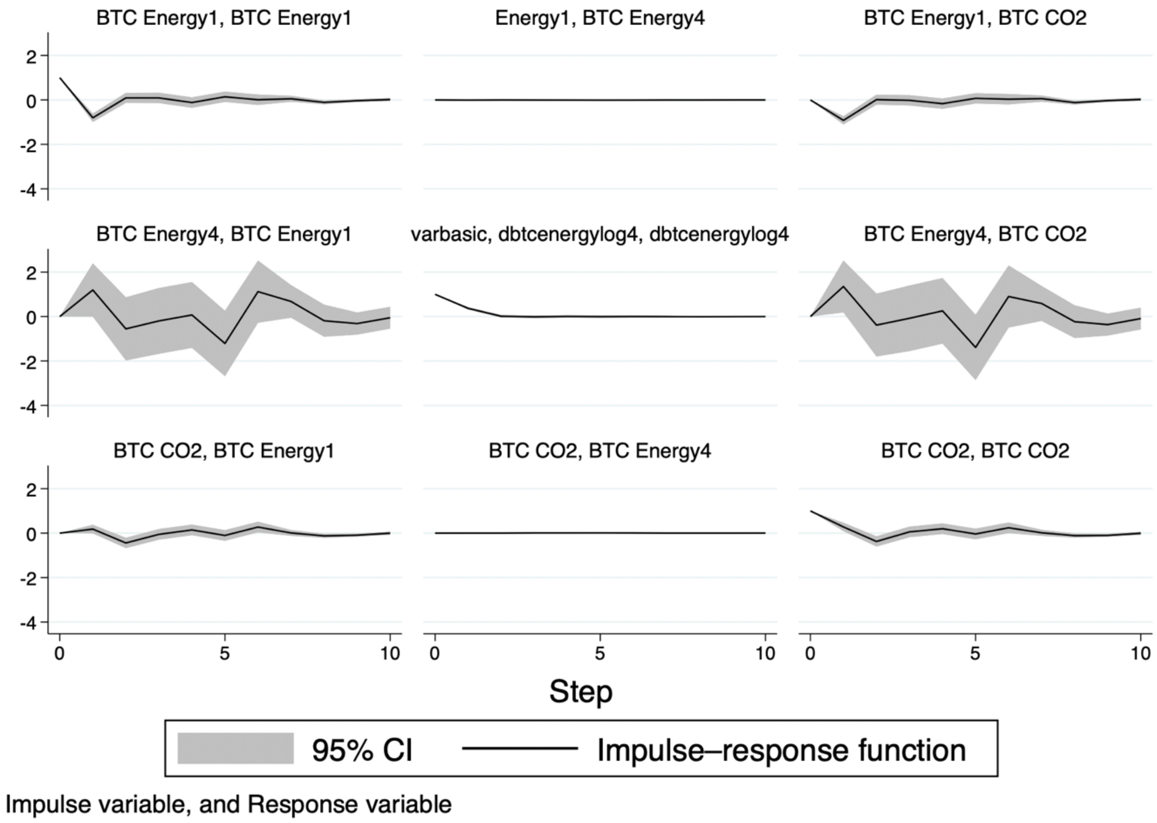

Figure 5 Results from IRF for Bitcoin

From those coefficients, I calculated the Granger causality. The coefficients for Bitcoin's Granger causality are displayed in Table 11. Focusing on explaining the carbon emissions, all the variables are relevant individually and with a 99.9% confidence interval. For explaining the first polynomial of energy consumption, all the variables are found to be relevant separately and together. Thus, there is a two-way Granger causation, implying selection bias because of endogeneity from the variables. Endogeneity occurs because of simultaneity (Antonakis et al., 2014). As with the Ethereum blockchain, there was insufficient public data to develop a good IV. Thus, I am not able to proceed with 2SLS. The correlation matrix for IV against the residuals is displayed in appendix A, Table 32A.

Table 11 Granger Causality Wald Tests for Bitcoin's carbon emissions

Y-variable	Exclude	DF	P-value
CO2 BTC	Energy BTC first polynomial	1	0
CO2 BTC	Energy BTC fourth polynomial	1	0
CO2 BTC	ALL	2	0
Energy BTC first polynomial	CO2 BTC	1	0
Energy BTC first polynomial	Energy BTC fourth polynomial	1	0
Energy BTC first polynomial	ALL	2	0
Energy BTC fourth polynomial	CO2 BTC	1	0.463
Energy BTC fourth polynomial	Energy BTC first polynomial	1	0.02
Energy BTC fourth polynomial	ALL	2	0

Note. BTC stands for Bitcoin. CO2 represents carbon emission. Energy relates to energy consumption. DF means degrees of freedom.

The residuals from the VAR model are stationary, and the variables present cointegration up to the first rank (appendix A Table 33A and 34A). Cointegration suggests that although the variables display short-term fluctuations, they move together in the long run, sharing a common underlying trend append. VECM models the long and short-term relationships between the variables. The VECM for Bitcoin is shown in expression [14]. The R-squared from Bitcoin's CO2 is 0.769, signalling that the model explains 76% of its variance. The long-term relationship between CO2 and energy in its first polynomial has a negative relationship, while CO2 and energy in the fourth polynomial has a negative relationship.

[14] VECM for Bitcoin Full Sample

Variables	CO2 BTC	Energy1 BTC	Energy 4 BTC
R-squared	0.769	0.761	0.275

$$Yt = \begin{bmatrix} -1.236 \\ -1.206 \\ 0.003 \end{bmatrix} CE_{t-1} + \begin{bmatrix} 0.508 & 2.264 & 0.634 \\ 1.256 & 1.433 & 0.544 \\ 0.002 & 0.002 & -0.473 \end{bmatrix} LD_1 + \begin{bmatrix} 0.220 & 1.081 & 0.297 \\ 0.785 & 1.179 & 0.101 \\ 0.000 & 0.001 & -0.490 \end{bmatrix} LD_2$$

$$+ \begin{bmatrix} -0.133 & 1.245 & 0.585 \\ 0.493 & 0.838 & 0.354 \\ 0.002 & -0.001 & -0.350 \end{bmatrix} LD_3 + \begin{bmatrix} -0.068 & 0.691 & 0.738 \\ 0.293 & 0.430 & 0.369 \\ 0.001 & 0.000 & -0.236 \end{bmatrix} LD_4$$

$$+ \begin{bmatrix} -0.078 & 0.301 & -0.687 \\ -0.011 & 0.213 & -0.768 \\ 0.003 & -0.003 & -0.150 \end{bmatrix} LD_5 + \begin{bmatrix} 0.000 \\ 0.000 \\ 0.000 \end{bmatrix}$$

Variable	(1) CE
CO2 BTC	1
Energy BTC 1st Polynomial	0.256***
	(-0.108)
Energy ETH 4^{TH} Polynomial	-0.354
	(0.884)
Constant	0.007

5.2.3 Hypothesis 2 Ecosystem

With both models defined, I modelled the CO2 for the ecosystem. This is done to test the second hypothesis. The expression from the ecosystem model [2] has been defined in the methodology. Following this, the linear model in Table 12 was calculated using OLS and NW standard errors. With this approach, only the Ethereum blockchain is significant. Both regressions can be found in Table 12.

Table 12 Model for the ecosystem's carbon emissions

Variable	(1) OLS	(2) NW
CO2 BTC	0.050	0.050
	(0.066)	(0.058)
CO2 ETH	0.823^{***}	$0. 823^{***}$
	(0.078)	(0.123)
Constant	-0.000	-0.000
	(0.008)	(0.005)
Observations	1289	1289
R^2	0.083	

Note. Standard errors are in parentheses. The OLS column stands for the model using the Ordinary Least Square approach. The NW column represents the model applying the Newey-West standard error. ETH stands for Ethereum, and BTC for Bitcoin. CO2 represents carbon emission. Energy relates to energy consumption. $^{*}p < 0.05$. $^{**}p < 0.01$. $^{***}p < 0.001$.

To improve the understanding of how the variables interact, I found the optimum number of lags by minimising the BIC (6 lags) (Table 35A, appendix A). Given the optimum number of lags, I modelled the VAR for the ecosystem, which is shown in expression [15]. The various R-squared appear statistically significant but are not a perfect fit to the endogenous variables' variance (only CO2 BTC is moderate). Unlike the OLS model, Bitcoin has a negative effect on the ecosystem's carbon footprint, while Ethereum displays a positive impact.

[15] VAR Ecosystem

Variables	CO2 Ecosystem	CO2 BTC	CO2 ETH
R-squared	0.183	0.435	0.18

$$Y_t = \begin{bmatrix} -0.037 & -0.391 & 1.342 \\ -0.010 & -0.671 & 0.108 \\ 0.035 & -0.090 & -0.214 \end{bmatrix} Y_{t-1} + \begin{bmatrix} -0.092 & -0.263 & 0.360 \\ -0.027 & -0.750 & 0.024 \\ 0.033 & -0.145 & -0.266 \end{bmatrix} Y_{t-2}$$

$$+ \begin{bmatrix} -0.032 & -0.160 & 0.501 \\ -0.006 & -0.666 & 0.052 \\ -0.012 & -0.156 & -0.085 \end{bmatrix} Y_{t-3} + \begin{bmatrix} 0.012 & -0.183 & 0.073 \\ 0.004 & -0.635 & 0.028 \\ -0.003 & -0.126 & -0.096 \end{bmatrix} Y_{t-4}$$

$$+ \begin{bmatrix} 0.008 & -0.041 & 0.126 \\ -0.015 & -0.529 & 0.008 \\ -0.026 & -0.103 & -0.091 \end{bmatrix} Y_{t-5} + \begin{bmatrix} 0.060 & -0.041 & 0.038 \\ -0.016 & -0.219 & 0.039 \\ -0.009 & -0.004 & -0.024 \end{bmatrix} Y_{t-6} + \begin{bmatrix} -0.035 \\ -0.007 \\ -0.031 \end{bmatrix}$$

Based on the ecosystem's VAR model, despite its low to moderate fit, the impulse reaction analysis was performed (Figure 6). It shows a clear reaction from the ecosystem's carbon footprint when both DLTs generate an impulse. However, impulses from the ecosystem do not pose a significant reaction towards the individual's DLTs. Next to that, Bitcoin is affected by Ethereum but not the other way around.

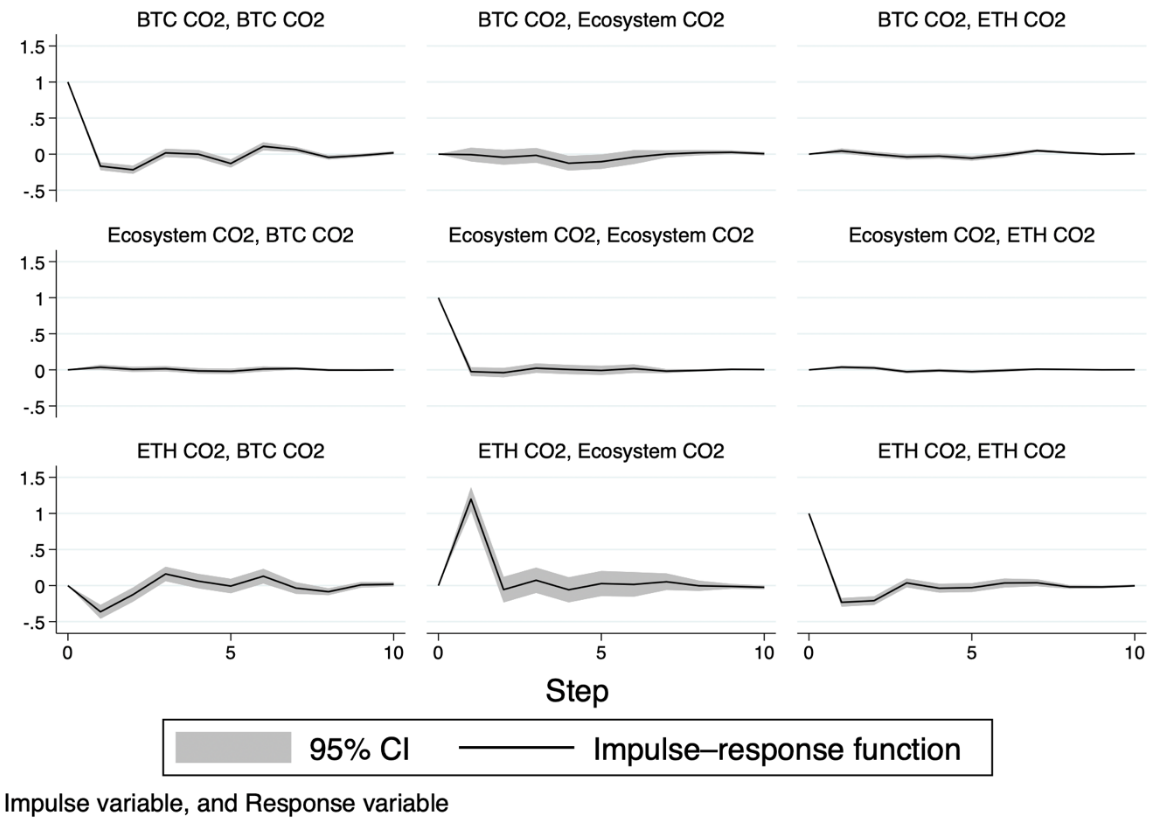

Figure 6 Results from IRF for the ecosystem

Upon estimating the Granger causality for the proposed VAR, the results reveal important relationships among the variables (see Table 13). Firstly, there is strong evidence of Granger causality between the carbon emissions of Ethereum, Bitcoin, and the ecosystem, as indicated by the p-values of 0.000. This suggests that changes in the ecosystem's CO2 emissions can be used to predict and explain variations in emissions from Bitcoin and Ethereum. Secondly, the Granger causality test indicates a significant relationship between Bitcoin and Ethereum, with a p-value of 0.021. This implies that fluctuations in Bitcoin may have an impact on Ethereum and vice versa. Additionally, Ethereum demonstrates strong Granger causality with the carbon footprint of the ecosystem and Bitcoin, further emphasizing the interconnectedness of the variables. The results of two-way Granger causality imply simultaneity, which is expected given that the ecosystem variable is generated by the summation of the original values of Ethereum and Bitcoin. When trying to derive an IV, all variables presented a correlation with the error term. Thus, no IV was found (see Table 36A, appendix A).

Table 13 Granger causality for the ecosystem carbon emissions

Y-variable	Exclude	DF	P-value
CO2 Ecosystem	CO2 BTC	1	0.000
CO2 Ecosystem	CO2 ETH	1	0.000
CO2 Ecosystem	ALL	2	0.000
CO2 BTC	CO2 Ecosystem	1	0.07
CO2 BTC	CO2 ETH	1	0.021
CO2 BTC	ALL	2	0.012
CO2 ETH	CO2 Ecosystem	1	0.000
CO2 ETH	CO2 BTC	1	0.000
CO2 ETH	ALL	2	0.000

Note. ETH stands for Ethereum and BTC for Bitcoin. CO2 represents carbon emission. Energy relates to energy consumption. DF means degrees of freedom.

As with the individual blockchains, the stationarity of the residuals from the ecosystem was tested. The residuals were stationary for the DF and PP test; however, the KPSS test found the non-stationary (Table 37A, appendix A). Despite the divergence of the data, I continued to test for cointegration. The Johansen-Juselius test found cointegration within the variables (Table 38A, appendix A). Therefore, I calculated the VECM for the ecosystem. The VECM model is defined in expression [16]. The R-squared for the ecosystem's CO2 is just below 0.5, indicating moderate explanatory power. Nevertheless, there is a large improvement in the OLS R-squared, while both Bitcoin's and Ethereum's carbon emissions have a high R-squared. A positive relationship exists between the ecosystem's emissions and both blockchain's individual emissions.

[16] VECM Ecosystem

Variable	CO2 Ecosystem	CO2 BTC	CO2 ETH
R-squared	0.466	0.77	0.505

$$\mathrm{Yt} = \begin{bmatrix} -0.005 \\ -0.049 \\ -0.010 \end{bmatrix} \mathrm{CE}_{t-1} + \begin{bmatrix} -0.859 & 0.114 & 1.214 \\ 0.037 & 2.806 & -0.157 \\ 0.023 & 0.814 & -0.984 \end{bmatrix} \mathrm{LD}_1 + \begin{bmatrix} -0.767 & -0.082 & 1.213 \\ 0.008 & 2.055 & -0.133 \\ 0.032 & 0.601 & -0914 \end{bmatrix} \mathrm{LD}_2 +$$

$$\begin{bmatrix} -0.592 & -0.082 & 1.255 \\ 0.001 & 1.388 & -0.080 \\ -0.008 & 0.376 & -0.633 \end{bmatrix} \mathrm{LD}_3 + \begin{bmatrix} -0.378 & -0.180 & 0.760 \\ 0.003 & 0.751 & -0.050 \\ -0.019 & 0.198 & -0.394 \end{bmatrix} \mathrm{LD}_4$$

$$+ \begin{bmatrix} -0.189 & -0.094 & 0.382 \\ -0.014 & 0.220 & -0.040 \\ -0.042 & 0.038 & -0.192 \end{bmatrix} \mathrm{LD}_5 + \begin{bmatrix} -0.001 \\ -0.000 \\ -0.000 \end{bmatrix}$$

Variable	(1) CE
CO2 Ecosystem	1
CO2 BTC	-8.007*** (-1.360)
CO2 ETH	-46.34*** (1.494)
Constant	0.574

Using the VECM, the log differences of the carbon emissions from the ecosystem per transaction were calculated. The summary statistics of the estimated variable can be found in Table 14.

Table 14 Descriptive statistics CO2 estimates

Variable	Mean	Median	*SD*	Max	Min	Skewness	Kurtosis	Observations
CO2 Estimate Ecosystem	0.009	0.008	0.085	1.244	-0.630	1.751	41.915	1290.000
CO2 Estimate BTC	-0.001	-0.009	0.120	0.586	-0.500	0.389	3.792	1290.000
CO2 Estimate ETH	0.018	0.018	0.102	1.519	-0.753	1.858	45.205	1290.000

Note. BTC stands for Bitcoin, ETH refers to Ethereum. S.D. means standard deviation, and CO2 represents carbon emission. All the values are calculated with log differences.

All the variables have a small difference between the mean and median and are close to zero. They also have a standard deviation close to one. This together implies a normal distribution for the values. The skewness of Bitcoin's estimated carbon emission is close to 0, which follows a normal distribution. But for the ecosystem and Ethereum, the skewness is closer to 2, implying skewness to the right. Bitcoin's Kurtosis again follows a normal distribution, while Ethereum and the ecosystem present a highly positive one, exhibiting leptokurtic behaviour. A plot with the estimated carbon emissions per transaction of Ethereum, Bitcoin and the ecosystem can be found in Figure 10B - 12B in appendix B.

To further inspect if the model has explanatory power, a two-sample t-test was implemented. The test compared the average estimated carbon emission per transaction of the ecosystem against the actual one. The t-test did not provide enough evidence to reject the hypothesis that the estimated CO2 equals the actual one. This implied that the proposed model does have some explanatory power. A summary of the test is shown in Table 15 below.

Table 15 Two-Sample T-test for estimated CO2 against actual values

Null Hypothesis	P-value
CO2 Estimate < Actual CO2	0.479
CO2 Estimate = Actual CO2	0.959
CO2 Estimate > Actual CO2	0.520

Note. CO2 relates to carbon emission.

5.3 Hypothesis 3

After building the models, structural breaks were checked according to the methodology. The breaks were tested in two ways, using energy consumption per transaction as an explanatory variable and applying the models proposed in the second hypothesis. The Chow break test was performed for the day that China banned cryptocurrencies and the day Ethereum changed its consensus approach. The p-values are displayed in Tables 16 and 17. For Ethereum, a break in trend was found in both regressions, following expectations. Results show that only the change in consensus created a structural break with a 95% confidence interval. This implies that the Chinese legislation, despite its scope, did not change Bitcoin's carbon footprint trend. However, Chinese regulation would be a significant structural break if considering expression [7] at a 90% confidence interval (see methodology). Moreover, the regulatory break was not found when considering only electricity consumption in the proposed model from hypothesis 2.

Table 16 Results Chow-break considering the expression [4]

Null Hypothesis	P-value
No Structural Break on ETH	0.000
No Structural Break on BTC	0.076

Note. BTC stands for Bitcoin, and ETH for Ethereum.

Table 17 Results Chow-break considering proposed models [9 and 10]

Null Hypothesis	P-value
No Structural Break on ETH	0.000
No Structural Break on BTC	0.111

Note. BTC stands for Bitcoin, and ETH for Ethereum.

The check made using the QLR did not find any extra relevant structural breaks. Because a significant break was found for Ethereum, hypotheses 1 and 2 must be re-examined for Ethereum and the ecosystem, but not for Bitcoin individually.

5.3.1 Reassessment of hypothesis 1

Before reassessing the first hypothesis, I tested all the variables for unit roots considering the structural break. The same findings (Tables 2A-7A in appendix A) were obtained when not pondering structural breaks, meaning all variables were non-stationary in absolute terms. But when transformed into its log differences, the variables became stationary.

When re-examining the first hypothesis with the structural breaks, there was not a significant difference between the average log difference of carbon emission per transaction of both blockchains, even after considering the first structural break. Yet, when considering the absolute values, a statistical difference between the blockchain was found. Their p values can be found in Tables 18 until 21.

Table 18 Two-sample T-Test before the break

Null Hypothesis	P-value
CO2 BTC < CO2 ETH	0.000
CO2 BTC = CO2 ETH	0.000
CO2 BTC > CO2 ETH	1.000

Note. BTC stands for Bitcoin, and ETH for Ethereum. CO2 relates to carbon emission. Values are in absolute terms (levels).

Table 19 Two-sample T-Test after the break

Null Hypothesis	P-value
CO2 BTC < CO2 ETH	0.006
CO2 BTC = CO2 ETH	0.011
CO2 BTC > CO2 ETH	0.995

Note. BTC stands for Bitcoin, and ETH for Ethereum. CO2 relates to carbon emission. Values are in absolute terms (levels).

Table 20 Two-sample T-Test before the break in logs

Null Hypothesis	P-value
CO2 BTC < CO2 ETH	0.523
CO2 BTC = CO2 ETH	0.952
CO2 BTC > CO2 ETH	0.476

Note. BTC stands for Bitcoin, and ETH for Ethereum. CO2 relates to carbon emission. Values are in log differences (changes).

Table 21 Two-sample T-Test after the break in logs

Null Hypothesis	P-value
CO2 BTC < CO2 ETH	0.848
CO2 BTC = CO2 ETH	0.303
CO2 BTC > CO2 ETH	0.151

Note. BTC stands for Bitcoin, and ETH for Ethereum. CO2 relates to carbon emission. Values are in log differences (changes).

5.3.2 Hypothesis 2 before the break

Focusing on the regression for Ethereum's carbon emissions before the break, energy is relevant in its first and fourth polynomials. This is considering OLS and NW standard errors. Differently from before, the second and third polynomials are irrelevant. Additionally, the first and second polynomial of the hashing power is found to be relevant and positive. The price in its first polynomial was also found to be relevant. No other variable was found to be relevant. The entire bottom-up approach can be found in the appendix. A total of 24 regressions were tested when deriving the log difference of carbon emission per transaction for the Ethereum blockchain before the structural break (17A - 20A in appendix A). The final model and its coefficients can be found in Table 22.

Table 22 Model for Ethereum before the structural break

Variable	(1) OLS	(2) NW
Energy ETH	0.897*** (0.009)	0.897*** (0.011)
Energy ETH 4th Polynomial	-1.220*** (0.214)	-1.220* (0.493)
Hashing power ETH	0.959*** (0.035)	0.959*** (0.038)
Hashing power ETH 2nd Polynomial	-2.317* (1.121)	-2.317* (1.145)
Price ETH	-0.353*** (0.015)	-0.353*** (0.035)
Constant	0.003** (0.001)	0.003** (0.001)
Observations	1198	1198
R^2	0.901	

Note. Standard errors are in parentheses. The OLS column stands for the model using the Ordinary Least Square approach. The NW column represents the model applying Newey-West standard error. ETH stands for Ethereum. CO2 represents carbon emission. Energy relates to energy consumption. $^{*}p < 0.05$. $^{**}p < 0.01$. $^{***}p < 0.001$.

What leaps out of the page is that the coefficient for the fourth polynomial is negative and bigger than one, while the first polynomial is positive and smaller than one. This implies that, for numbers smaller than one, energy will cause a positive effect on carbon emissions, while large changes in energy have a negative impact. This does go against what is suggested in the literature. It is expected that energy affects carbon emissions in a positive manner only. For the hashing power, the same pattern was found. This also contradicts the literature. Further, price is also found to have a negative coefficient. This suggests that when there is a price increase, there is less carbon used in a transaction. This goes against the idea that a higher price of the native token will result in a higher carbon emission because there will be more competition. According to Ethereum's model, before the break, price, and therefore competition, drive down carbon emissions. Additionally, the constants were found relevant and positive. This implies that even if all other variables are zero, some underlying variable still generates carbon emission.

As was done before, the Granger causality using the VAR approach considering BIC was calculated. In this case, the optimum number of lags is one (Table 39A in appendix A). The VAR derivation resulted in expression [17]. Based on the VAR model, the various R-squared are close to zero. The R-squared indicates the proportion of the variance in each equation explained by the exogenous variables and suggests, in this case, that the variables have limited explanatory power. Only when considering the carbon footprint equation a similar relationship between the variables can be found in comparison to the NW model. As an example, the relationship direction (positive or negative) is the same. Different from the previous model, energy in its fourth polynomial became the largest coefficient.

[17] VAR Ethereum before the break

Variables	CO2 ETH	Energy1 ETH	Energy4 ETH	Hashing1 ETH	Hashing4 ETH	Price ETH
R-squared	0.056	0.047	0.013	0.13	0.033	0.004

$$Y_t = \begin{bmatrix} -0.446 & 0.250 & -2.403 & 0.220 & -0.741 & -0.077 \\ -0.406 & 0.205 & -1.964 & 0.575 & -0.512 & -0.063 \\ -0.006 & 0.008 & 0.058 & 0.012 & -0.086 & -0.002 \\ -0.060 & 0.057 & -0.056 & -0.292 & 0.225 & 0.009 \\ 0.001 & -0.001 & 0.000 & 0.002 & 0.138 & 0.001 \\ -0.087 & 0.068 & 0.387 & 0.148 & -0.879 & -0.048 \end{bmatrix} Y_{t-1} + \begin{bmatrix} 0.001 \\ -0.000 \\ 0.000 \\ -0.002 \\ 0.000 \\ -0.002 \end{bmatrix}$$

The IRF (see Figure 7) shows that no variables present a long-term reaction from the impulse of other variables. The hashing power in its second polynomial is the variable which creates the largest reaction from the carbon footprint. This goes against what was expected from the VAR model because the variable with the highest coefficient was energy consumption in its fourth polynomial. In contrast, the small reactions from the impulses were expected given the small R-squared for the vectors. An overview of all the impulses and responses is shown in Figure 7.

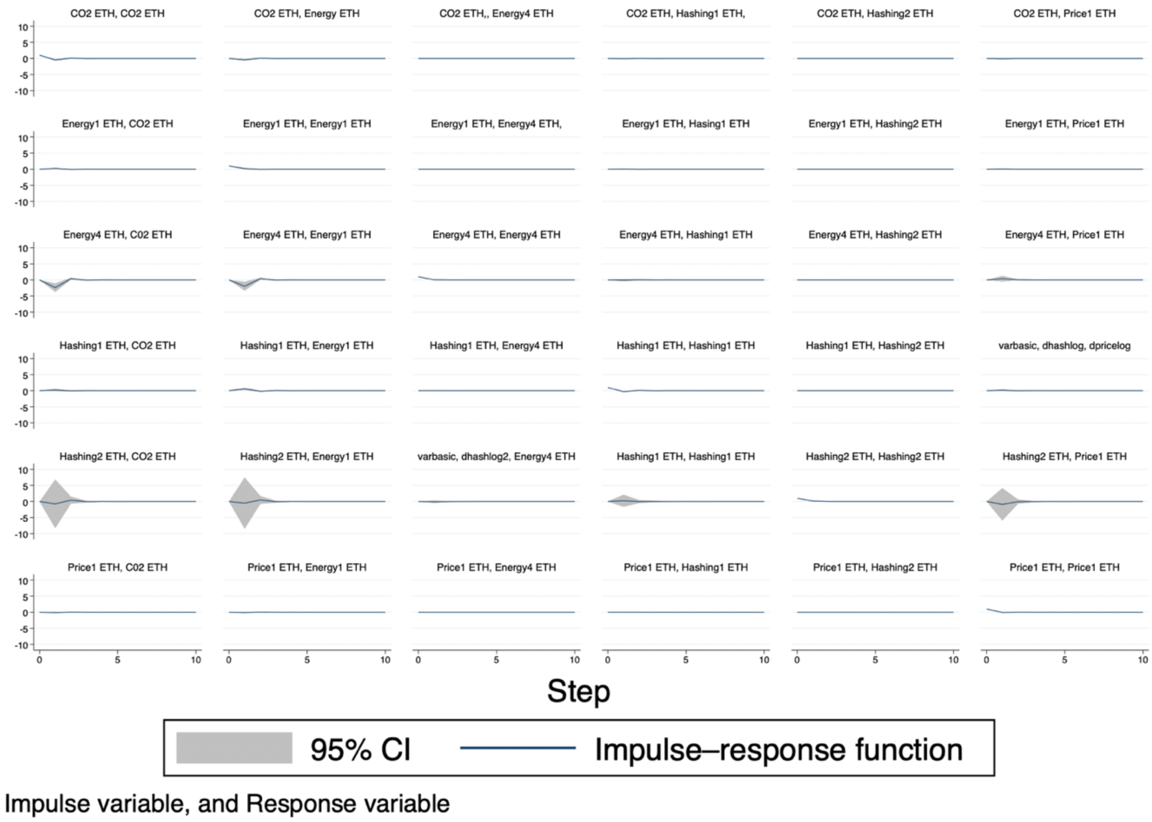

Figure 7 Results from IRF for Ethereum before the break

This is to check if the model has some prediction power which could relate to explanatory as well. The p-values for all the individual factors and their clustered can be found in Table 23.

Table 23 Granger Causality Wald Tests for Ethereum before the structural break

Y-variable	Exclude	DF	P-value
CO2 ETH	Energy ETH first polynomial	1	0.004
CO2 ETH	Energy ETH fourth polynomial	1	0.000
CO2 ETH	Hashing power ETH first polynomial	1	0.123
CO2 ETH	Hashing power ETH second polynomial	1	0.840
CO2 ETH	Price ETH first polynomial	1	0.161
CO2 ETH	All	5	0.003
Energy ETH first polynomial	CO2 ETH	1	0.000
Energy ETH first polynomial	Energy ETH fourth polynomial	1	0.005
Energy ETH first polynomial	Hashing power ETH first polynomial	1	0.000
Energy ETH first polynomial	Hashing power ETH second polynomial	1	0.900
Energy ETH first polynomial	Price ETH first polynomial	1	0.2817
Energy ETH first polynomial	All	5	0.000
Energy ETH fourth polynomial	CO2 ETH	1	0.164

Energy ETH fourth polynomial	Energy ETH first polynomial	1	0.043
Energy ETH fourth polynomial	Hashing power ETH first polynomial	1	0.075
Energy ETH fourth polynomial	Hashing power ETH second polynomial	1	0.627
Energy ETH fourth polynomial	Price ETH first polynomial	1	0.498
Energy ETH fourth polynomial	All	5	0.121
Hashing power ETH first polynomial	CO2 ETH	1	0.008
Hashing power ETH first polynomial	Energy ETH first polynomial	1	0.008
Hashing power ETH first polynomial	Energy ETH fourth polynomial	1	0.735
Hashing power ETH first polynomial	Hashing power ETH second polynomial	1	0.841
Hashing power ETH first polynomial	Price ETH first polynomial	1	0.527
Hashing power ETH first polynomial	All	5	0.011
Hashing power ETH second polynomial	CO2 ETH	1	0.105
Hashing power ETH second polynomial	Energy ETH first polynomial	1	0.384
Hashing power ETH second polynomial	Energy ETH fourth polynomial	1	0.992
Hashing power ETH second polynomial	Hashing power ETH first polynomial	1	0.115
Hashing power ETH second polynomial	Price ETH first polynomial	1	0.105
Hashing power ETH second polynomial	All	5	0.001
Price ETH first polynomial	CO2 ETH	1	0.155
Price ETH first polynomial	Energy ETH first polynomial	1	0.241
Price ETH first polynomial	Energy ETH fourth polynomial	1	0.380
Price ETH first polynomial	Hashing power ETH first polynomial	1	0.120
Price ETH first polynomial	Hashing power ETH second polynomial	1	0.732
Price ETH first polynomial	All	5	0.499

Note. ETH stands for Ethereum. CO2 represents carbon emission. Energy relates to energy consumption. DF means degrees of freedom.

The Granger causality test reveals significant relationships among variables in the Ethereum ecosystem. The tests indicate that carbon emissions and energy consumption have a statistically significant causal relationship. Additionally, carbon emissions are causally related to lagged energy consumption. Moreover, there is a significant causal relationship between energy consumption and hashing power, while hashing power is also causally related to carbon emissions. Furthermore, a lagged version of hashing power shows a significant causal relationship with carbon emissions. These findings shed light on the underlying simultaneity between the variables. Such behaviour occurs because of endogeneity, which can be the result of omitted variables bias or measurement errors (Antonakis et al., 2014).

Because of the Granger dual causality, possible IVs were inspected. Unfortunately, all variables were correlated with the residuals of the model (40A in the appendix A). Thus, I am not able to process with the 2SLS approach. However, there is still the possibility of cointegration between the variables. When checking if the residuals from the model are stationary, the same was found to be true (41A in the appendix A). This allows me to test for cointegration with the Johansen-Juselius cointegration rank test. Overall, the results suggest that the variables in the dataset are interconnected and have long-run and short-run relationships. The Johansen-Juselius cointegration rank test results are in appendix A Table 42A.

Because cointegration was found in the VAR model, I can employ VECM. The VECM for Ethereum before the break is defined in expression [18]. The VECM resulted in equations with mixed r-squares, where some are extremely close to 0, whereas others are above 0.5. Further, the cointegration equation, which represents the long-run relationship of the exogenous variables against the endogenous. In this case, energy, price and hashing power in its second polynomial have a negative relation with carbon footprint, while hashing rate and the constant have a positive one. The short-term coefficients are all small and, in their majority, also present a negative relationship with carbon emissions.

[18] VECM Ethereum before the break

Variables	CO2 ETH	Energy1 ETH	Energy4 ETH	Hashing1 ETH	Hashing4 ETH	Price ETH
R-squared	0.029	0.008	0.002	0.663	0.037	0.008

$$Y_t=\begin{bmatrix}-0.067\\0.035\\0.001\\-0.087\\0.000\\0.020\end{bmatrix}CE_{t-1}+\begin{bmatrix}0.000\\0.000\\0.000\\0.000\\0.000\\0.000\end{bmatrix}$$

Variable	(1) CE
CO2 ETH	1
Energy1 ETH	-1.026***
	(0.559)
Energy4 ETH	-0.089
	(1.331)
Hashing1 ETH	9.693***
	(0.218)
Hashing4 ETH	-43.3621***
	(7.033)
Price ETH	-0.289***
	(0.092)
Constant	0.031

5.3.3 Hypothesis 2 after the break

I considered 16 models to derive the Ethereum carbon emission after the break. For the model after the structural break, only energy consumption in its first polynomial was found to be relevant. The regression has enough evidence to reject the null hypothesis of constant variance in the error term. Further, it was found that a serial auto relation exists up to its first lag. Further, the first polynomial was only found to be relevant when using NW standard errors. The regression can be found in Table 24.

Table 24 Model for Ethereum after the structural break

Variable	(1) OLS	(2) NW
Energy ETH	0.233	0.233**
	(0.367)	(0.069)
Constant	0.241	0.241
	(0.256)	(0.253)
Observations	90	90
R^2	0.005	

Note. Standard errors are in parentheses. The OLS column stands for the model using the Ordinary Least Square approach. The NW column represents the model applying the Newey-West standard error. ETH stands for Ethereum. CO2 represents carbon emission. Energy relates to energy consumption. $^{*}p < 0.05$. $^{**}p < 0.01$. $^{***}p < 0.001$.

The small R-squared value in the model for Ethereum after the break leaps out of the page. This can be the result of the small data sample. Also, the only driver for a carbon footprint is energy which has a positive coefficient, which follows the literature expectations. The bottom-up approach for the log differences from carbon emissions from Ethereum after the break can be found in Tables 21A - 24A, appendix A.

Before I estimated the VAR model, I calculated the optimum number of lags considering the BIC, which is three (Table 43A, appendix A). The VAR model is defined in expression [19]. The VAR model presents a large improvement in the R-squared. The VAR indicates a negative correlation with Ethereum's carbon emissions with its lags and a positive one with Ethereum's energy consumption in its first polynomial.

[19] VAR Ethereum after the break

Variables	CO2 ETH	Energy1 ETH
R-squared	0.914	0.358

$$Y_t = \begin{bmatrix} -1.257 & 1.677 \\ 0.367 & 0.890 \end{bmatrix} Y_{t-1} + \begin{bmatrix} -0.555 & 4.023 \\ 0.095 & -1.383 \end{bmatrix} Y_{t-2} + \begin{bmatrix} -0.038 & 2.063 \\ -0.014 & -0.311 \end{bmatrix} Y_{t-3} + \begin{bmatrix} -0.010 \\ 0.057 \end{bmatrix}$$

From the VAR model, I estimated the IRF between the variables. Results from the IRF are in Figure 8. The carbon emissions of Ethereum are highly affected by itself and its energy consumption. On the other hand, energy consumption appears not to have large responses when faced with an impulse.

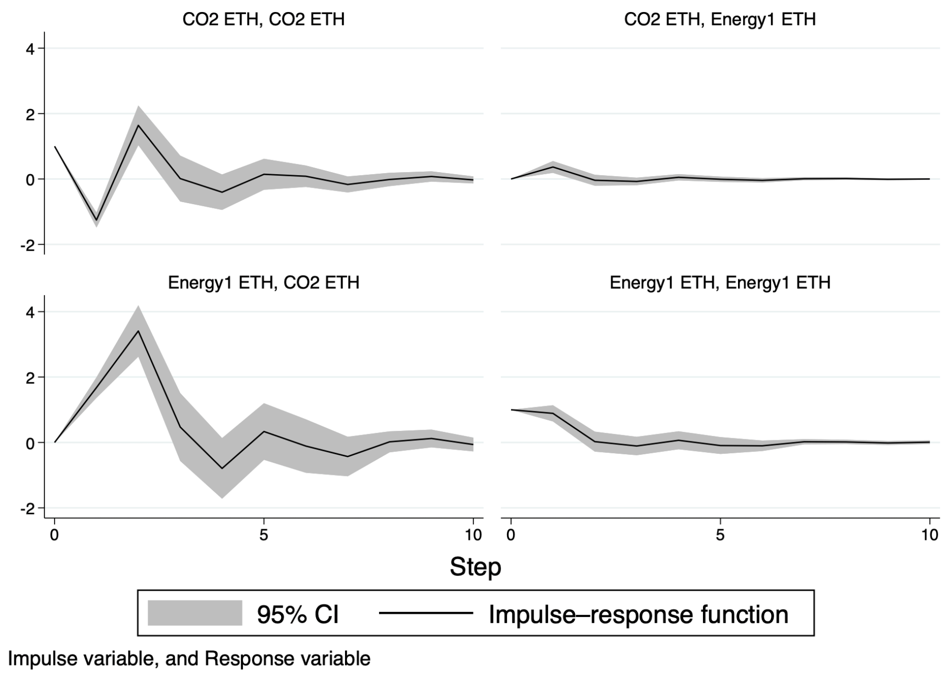

Figure 8 Results from IRF for Ethereum after the structural break

As conducted earlier, the Granger causality test was performed. The results (see Table 25) present the p-values obtained from the VAR model and indicate a significant two-way Granger causality relationship between Ethereum's carbon emissions and Ethereum's energy consumption. This means that changes in Ethereum's energy consumption can help predict and explain variations in its carbon footprint and vice versa. Because of this bias, I applied the 2SLS approach. Unfortunately, as before, no relevant IV was found because all the variables had a correlation with the VAR residuals (Table 44A, appendix A).

Table 25 Granger Causality Wald Tests for Ethereum after the structural break

Y-variable	Exclude	DF	P-value
CO2 ETH	Energy ETH	1	0.000
CO2 ETH	all	1	0.000
Energy ETH	CO2 ETH	1	0.000
Energy ETH	all	1	0.000

Note. ETH stands for Ethereum. CO2 represents carbon emission. Energy relates to energy consumption. DF means degrees of freedom.

Given the two-way Granger causality, I tested if the residuals were stationary. The same results as for Ethereum before the break were found, meaning that residuals for the DF and PP tests were found stationary, whereas for KPSS non-stationary (Table 45A, appendix A). Once again, despite the divergence in the results, I continued to test for cointegration. The Johansen-Juselius cointegration rank test found cointegration within the variable (Table 46A, appendix A). Therefore, I estimated the VECM. The VECM for Ethereum after the break is shown in the Expression [20].

[20] VECM Ethereum after break

Variables	CO2 ETH	Energy1 ETH
R-squared	0.953	0.328

$$Y_t=\begin{bmatrix}-2.904\\0.558\end{bmatrix}CE_{t-1}+\begin{bmatrix}0.627 & -6.040\\-0.149 & 1.600\end{bmatrix}LD_1+\begin{bmatrix}0.063 & -1.999\\-0.036 & 0.181\end{bmatrix}LD_2+\begin{bmatrix}-0.000\\-0.002\end{bmatrix}$$

Variable	(1) CE
CO2 ETH	1
Energy1 ETH	-1.026*** (0.055)
Constant	-0.007

The VECM model for Ethereum after the break has significant R-squares. Further, for the carbon footprint equation, the R-square is 0.95, which indicates a good fit within the model. However, the model for Ethereum's energy consumption in its first polynomial has a low R-squared. The cointegration equation indicates that Ethereum's carbon emissions and energy consumption have a long-term positive relationship with each other. For the carbon emissions equation, in contrast with the VAR model, the coefficients of energy consumption became negative, and the coefficients of the carbon emissions lags became positive.

5.3.4 Merging models

To derive the general ecosystem formula, the predicted values for the Ethereum blockchain before and after the structural break were merged by connecting both data samples on the break day. Subsequently, the new estimate of Ethereum's carbon emissions was regressed with the previously estimated emissions of Bitcoin. The result can be found in Table 26.

Table 26 Model for ecosystem considering the structural break

Variable	(1) OLS	(2) NW
CO2 Estimate BTC	0.088 (0.065)	0.088 (0.069)
CO2 Estimate ETH	0.767*** (0.073)	0.767*** (0.196)
Constant	-0.005 (0.008)	-0.005 (0.004)
Observations	1289	1289
R^2	0.089	

Note. Standard errors are in parentheses. The OLS column stands for the model using the Ordinary Least Square approach. The NW column represents the model applying the Newey-West standard error. ETH stands for Ethereum, and BTC for Bitcoin. CO2 represents carbon emission. $^{*}p < 0.05$. $^{**}p < 0.01$. $^{***}p < 0.001$.

For the ecosystem model considering structural breaks, only Ethereum proves relevant to explain the changes in the ecosystem's carbon emission per transaction. This result contradicts expectations, given that Bitcoin holds a larger market value and user base and enjoys greater public recognition as a DLT than Ethereum. Besides that, both heteroskedasticity and serial autocorrelation were found.

Following the methodology, I minimised the BIC to discern the optimum number of lags in the VAR model. The minimisation showed two lags as optimal (see appendix A, Table 47A). The VAR model is shown in expression [21]. The R-squares of all the equations in the VAR model are close to 0. This could indicate that other variables not included in the model may contribute to a more comprehensive understanding of the variables' relationship.

[21] VAR general mode for Ecosystem considering structural break

Variables	CO2 Ecosystem	CO2 BTC	CO2 ETH
R-squared	0.035	0.443	0.163

$$Y_t = \begin{bmatrix} 0.121 & 0.035 & -0.016 \\ -0.036 & -0.152 & 0.356 \\ -0.005 & 0.005 & -0.393 \end{bmatrix} Y_{t-1} + \begin{bmatrix} -0.038 & -0.025 & -0.131 \\ -0.002 & -0.249 & 0.158 \\ 0.001 & 0.002 & -0.134 \end{bmatrix} Y_{t-2} + \begin{bmatrix} 0.006 \\ 0.000 \\ -0.00 \end{bmatrix}$$

From the VAR model, I calculated the IRF. The IRF is displayed in Figure 9. From the IRF, it is possible to infer that carbon emission impulses generated by the ecosystem will create a permanent positive effect on itself. Further, the ecosystem does not affect the blockchains individually, whereas the blockchains do affect the ecosystem.

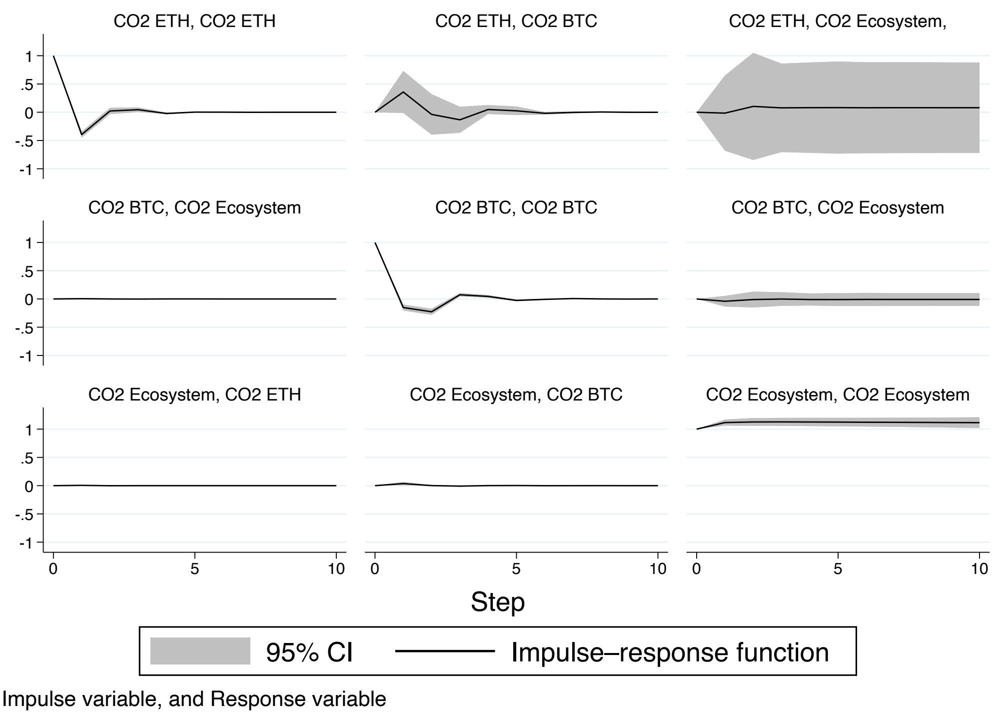

Figure 9 IFR for The ecosystem considering the structural break

Following the IRF analysis, the Granger causality test (Table 27) examined bidirectional causality between variables. Results indicate no significant unidirectional Granger causality within the models at a 95% confidence level. However, as anticipated, closer examination reveals nearly significant findings, with three p-values demonstrating 94% confidence levels. This highlights the interactive impact of DLT systems on individual blockchains.

Table 27 Granger Causality Wald Test for ecosystem considering the structural break

Y-variable	Exclude	DF	P-value
CO2 Ecosystem	CO2 BTC	2	0.630
CO2 Ecosystem	CO2 ETH	2	0.924
CO2 Ecosystem	ALL	4	0.900
CO2 BTC	CO2 Ecosystem	2	0.059
CO2 BTC	CO2 ETH	2	0.168
CO2 BTC	ALL	4	0.059
CO2 ETH	CO2 Ecosystem	2	0.054
CO2 ETH	CO2 BTC	2	0.389
CO2 ETH	ALL	4	0.121

Note. ETH stands for Ethereum. CO2 represents carbon emission. Energy relates to energy consumption. DF means degrees of freedom.

The VAR model residuals confirmed stationary, the same way it did for Ethereum pre- and post-break (see Table 48A). Based on this, cointegration was examined. The Johansen-Juselius test indicated first-rank cointegration (see Table 49A). Consequently, the VECM was computed to analyse the ecosystem's carbon footprint (expression [22]).

[22] VECM general model for ecosystem consideringstructural break

Variables	CO2 Ecosystem	CO2 BTC	CO2 ETH
R-squared	0.348	0.179	0.029

$$Y_t=\begin{bmatrix}0.001\\-0.007\\0.001\end{bmatrix}CE_{t-1}+\begin{bmatrix}-0.422 & 0.093 & 0.084\\-0.029 & 0.056 & -2.483\\-0.007 & -0.045 & -0.384\end{bmatrix}LD+\begin{bmatrix}0.006\\0.000\\0.000\end{bmatrix}$$

Variable	(1) CE
CO2 Ecosystem	1
CO2 BTC	139.004*** (4.622)
CO2 ETH	-713.248*** (35.643)
Constant	0.031

Results of the VECM model for the ecosystem's carbon footprint equation showed a moderate explanatory power (R-squared = 0.348), while Bitcoin and Ethereum equations exhibited a weak explanatory power (R-squared = 0.179, R-squared = 0.029). The cointegration equation confirmed a long-term relationship among the variables, with a positive correlation observed for Bitcoin's carbon footprint and a negative correlation for Ethereum's carbon footprint relative to the ecosystem's carbon footprint.

Using the VECM model, expected values for the ecosystem were estimated. Table 28 presents descriptive statistics of estimated carbon emissions from the Ethereum blockchain and the ecosystem, accounting for a structural break. Bitcoin's carbon emission estimate was excluded because it was not changed.

Table 28 Descriptive statistics of CO2 estimates for Ethereum and ecosystem

Variable	Mean	Median	*SD*	Max	Min	Skewness	Kurtosis	Observations
CO2 Estimate ETH	0.018	0.002	0.108	1.562	-0.436	3.140	39.731	1290
CO2 Estimate Ecosystem	0.009	-0.004	0.289	9.807	-0.388	30.511	1033.184	1289

Note. ETH refers to Ethereum. S.D. means standard deviation, and CO2 represents carbon emission. All the values are calculated with log differences.

Similar to the descriptive statistics from before the breaks, the mean and median are close to each other and zero. This indicates a normal distribution. Further, the standard deviation is close to one, suggesting a normal distribution. Both variables' skewness is larger than zero, therefore not following a normal distribution. The variables present skewness to the right. Their kurtosis is also considerably large, which represents leptokurtic behaviour. A plot with the estimated carbon footprint per transaction of Ethereum and the ecosystem considering the break can be found in Figures 13B and 14B, appendix B.

To further test the proposed model's explanatory power, I calculated the average log difference between the actual and predicted ecosystem carbon footprint. These test results can be found in Table 29. Based on the two-sample t-test, there is insufficient evidence to reject the hypothesis that the average log difference between estimated and actual carbon emissions of the ecosystem is not significant, suggesting some explanatory power from the proposed model.

Table 29 Two-sample T-Test actual and predicted ecosystem carbon emissions

Null hypothesis	P-value
CO2 Estimate Ecosystem < Actual CO2 Ecosystem	0.790
CO2 Estimate Ecosystem = Actual CO2 Ecosystem	0.419
CO2 Estimate Ecosystem > Actual CO2 Ecosystem	0.209

Note. CO2 relates to carbon emission.

Regarding the third hypothesis, no structural break was observed for Bitcoin during the period of focus at a 95% confidence level. However, with a 90% confidence level, Chinese bans on cryptocurrencies would have resulted in a structural break for Bitcoin. Now Ethereum experienced a structural break due to a soft fork relevant at a 05% confidence interval.

When reassessing the first hypothesis, a similar result from not considering the ecosystem was found. Thus, the carbon footprint of DLTs is different in absolute values but presents a non-different average change.

For the reassessment of the second hypothesis, Bitcoin showed relevance in the first and fourth polynomials of energy consumption across the entire sample. Pre-break Ethereum demonstrated relevance in energy consumption, hashing power, and price. Post-break, only the energy consumption in the first polynomial remained relevant.

When merging the model, in the simpler model, only Ethereum was considered relevant when explaining the carbon footprint. But when employing the more complex model, bitcoin also became relevant. The Granger causality test proves that the integration effect between the variables captured by the ecosystem Granger causes changes in the individual blockchains accounting for a 94% confidence interval.

All models exhibited cointegration, enabling the use of VECM. In most cases, the VECMs improved the modelled relationship, as became evident through increased R-squared values. No statistically significant differences were found when comparing all VECM-estimated values with actual values, verifying the proposed models' explanatory power.

5.4 Predictive power

The main goal of this section is to expand on the proposed model's explanatory power concerning the CO2 emissions of the ecosystem. To do so, I generated forecasts using the VECM models from hypothesis 3. The descriptive statistics of the forecasted variables can be found in Table 30. The forecasted values are made using the rolling window approach with 30 days for training the models. The mean, median and standard deviation for all forecasts follow a normal distribution. But when considering their skewness and kurtosis, a non-normal distribution is found. The ecosystem skewness is negative, and Bitcoin is large and negative, resulting in a distribution skewed to the left, whereas Ethereum has positive skewness. For kurtosis, all the variables present a large and positive value. The higher the kurtosis, the more concentrated the values around the mean are.

Table 30 Descriptive statistics of CO_2 estimates for ETH, BTC and ecosystem

Variable	Mean	Median	*SD*	Max	Min	Skewness	Kurtosis	Observations
CO2 Estimate ETH	0.392	-0.000	29.327	915.011	-466.384	21.411	822.668	1231.000
CO2 Estimate BTC	-0.064	-0.001	5.253	76.869	-166.711	-23.425	862.564	1231.000
CO2 Estimate Ecosystem	-0.001	-0.006	0.292	1.891	-3.246	-1.696	28.694	1231.000

Note. BTC stands for Bitcoin, ETH refers to Ethereum. S.D. means standard deviation, and CO2 represents carbon emission. All the values are calculated with log differences.

To inspect the explanatory power of the forecasts following the rolling window approach, a two-sample T-test was made against the actual values. There was not enough evidence to reject the null hypothesis of the average log difference of carbon emission from the forecast values being different from the actual ones considering a confidence level of 95%. The p-values of the tests are displayed in Table 50A, appendix A. This implies that the proposed forecast has some explanatory power.

To further assess the explanatory power of the models, a comparison against a benchmark was made. The benchmark composes of an autoregressive model taking into account only the first lag of the endogenous variable, as explained in the methodology expression [10]. The descriptive statistics of the benchmark are shown in Table 31 below.

Table 31 Descriptive statistics of CO_2 benchmarks for ETH, BTC and ecosystem

Variable	Mean	Median	*SD*	Max	Min	Skewness	Kurtosis	Observations
CO2 Benchmark ETH	0.010	-0.000	0.091	2.506	-0.167	18.090	472.591	1231.000
CO2 Benchmark BTC	0.020	-0.001	0.139	2.506	-0.197	8.364	103.908	1231.000
CO2 Benchmark Ecosystem	-0.002	-0.001	0.031	0.182	-0.178	-0.758	8.064	1231.000

Note. BTC stands for Bitcoin, ETH refers to Ethereum. S.D. means standard deviation, and CO2 represents carbon emission. The benchmark relates to expression [8]. All the values are calculated with log differences.

The variables' mean and median are close to 0, following a normal distribution. The benchmark for Bitcoin has a value close to zero for the standard deviation, which also follows a normal distribution, conversely from the other variables. Considering the skewness, the one for Bitcoin is large and positive, indicating skewness to the left. Ethereum displays the same behaviour. The skewness of the ecosystem is close to 0. The kurtosis from all of them is large and positive.

With the forecasts, I applied a DM test to search for differences between the RMSE of the benchmark and the predicted values. The p-values from the test can be found in Table 32.

Table 32 Results from the DM test

Null hypothesis	P-value
RMSE CO2 Ecosystem Estimate < RMSE Ecosystem CO2 Benchmark	0.925
RMSE CO2 Ecosystem Estimate = RMSE CO2 Ecosystem Benchmark	0.149
RMSE CO2 Ecosystem Estimate > RMSE CO2 Ecosystem Benchmark	0.074
RMSE CO2 ETH Estimate < RMSE CO2 ETH Benchmark	0.552
RMSE CO2 ETH Estimate = RMSE CO2 ETH Benchmark	0.897
RMSE CO2 ETH Estimate > RMSE CO2 ETH Benchmark	0.448
RMSE CO2 BTC Estimate < RMSE CO2 BTC Benchmark	1.000
RMSE CO2 BTC Estimate = RMSE CO2 BTC Benchmark	0.000
RMSE CO2 BTC Estimate > RMSE CO2 BTC Benchmark	0.000

Note. CO2 relates to carbon emission. BTC stands for Bitcoin, ETH refers to Ethereum. The benchmark relates to the expression [8]

According to the DM test, there is not enough evidence to reject the hypothesis that the models have an equal RMSE for Ethereum and the ecosystem's carbon footprint. This implies that the proposed model does have strong explanatory power. This also means that the proposed model is equal to the benchmark in any situation. So, despite not adding any extra information against the lag of the endogenous variable, it can still model relevant information. However, the DM test states that the model average is larger and different than the benchmark for Bitcoin. Thus, the proposed model for Bitcoin has a worse predictive power than the benchmark model.

CHAPTER 6 Conclusion

In this chapter, I will accept or reject the hypotheses derived from the literature, while discussing their implications. After, the research question "What are the main drivers of the carbon footprint from the public blockchain ecosystem and its externalities between 2019-2022?" will be answered. Lastly, the research limitations will be discussed, together with recommendations for future research.

6.1 Hypothesis 1

The first hypothesis tested whether blockchains differ in their average carbon emission per transaction over time. Previous research looked for differences between different consensus mechanisms and found relevant results (de Vries et al., 2022). Blockchains with the same consensus mechanism have dissimilarities in their protocol, network, and application layers. However, those dissimilarities are less evident than for blockchains with different consensus mechanisms. To expand on the literature, the first hypothesis was derived:

"There is a statistical difference between the average carbon emissions per transaction across different DLTs."

A Two-sample T-Test on the cleaned variables showed a difference in the average carbon emissions per transaction between different DLTs. This is shown in the results Table 6. There is not enough statistical evidence to reject hypothesis one. In other words, dissimilarities in the DLTs generate statistically different carbon footprints in absolute terms. This is a favourable discovery for the future development of DLTs because it means that there are multiple ways to decrease carbon footprint significantly.

For the Two-sample T-Test considering the log difference of the carbon emissions per transaction, there was insufficient evidence of differences in trends. So, despite their statistically different absolute values, their average daily change is not significantly different. This result signals the difficulty of maintaining a sustainable DLT.

6.2 Hypothesis 2

Hypothesis 1 found an absolute difference in carbon emissions across DLTs, but not in their daily changes. The second hypothesis investigated the main drivers for the daily change in the carbon emissions of the blockchain ecosystem. The second hypothesis is:

"*Energy consumption, hashing rate, e-waste generation, and prices of native tokens have a significant explanatory effect on the blockchain's ecosystem carbon footprint.*"

Considering the whole sample, only energy consumption was relevant for explaining the carbon emissions of the ecosystem. Based on this, the second hypothesis was rejected. This is understandable because the most common method for calculating carbon emissions is based solely on energy use. Further, if the energy consumption of the ecosystem were to be computed using a direct method, it would most likely be based on other variables, such as hashing power and the price of the native tokens.

The VAR models, which I used to study carbon emissions in more detail, present an improvement in the R-squares of the OLS equations. The IRF showed that impulse in the carbon emissions of the ecosystem and individual DLTs do have a short-term response, but that the responses die out in the long run. The Granger test revealed a two-way Granger causality between the variables. When performing the 2SLS, all variables utilized in this research were found to have a correlation with the residuals from the VAR models. Thus, it was impossible to generate a relevant IV. I recommend future research to expand on IVs.

Trying to improve the understanding of carbon emissions' externalities, the residuals of the VAR models were found stationary, for two out of three tests employed. Further, the variables for all individual models have cointegration considering the VARs. Consequently, I was able to apply VECM model. The VECM models also improved the explanatory of the relationship, in comparison to the VAR. The cointegration coefficients indicate that the ecosystem has a positive and significant long-term relationship within the variables. This implies that regulations, motivating miners and stakers to run their computers on renewable energy sources, are likely to affect the ecosystem's carbon emissions in the long run. On the private side, parties should prioritize energy consumption to maintain the ecosystem when creating new DLTs to minimize their carbon footprint.

6.3 Hypothesis 3

In the last years, DLTs are facing increasing regulations and updates. To some extent, those changes are triggered by sustainability concerns. Therefore, I derived the third hypothesis for testing if those legislative and technical changes affect the ecosystem's carbon emissions. The hypothesis is:

“There is at least one structural break in a blockchain carbon footprint per transaction”.

With a 5% error level, the Chinese government's change in legislation did not have enough data to support the hypothesis that there was a structural break, but it did so on a 10% level. With more data, relevant breaks could have been found. For Ethereum, the soft fork that changed its protocol was found to have caused a significant structural break. Consequently, there is insufficient evidence to reject the hypothesis. This means that updating DLT layers makes it possible to change the carbon emission trend. The impact of the structural break goes in line with what is expected. The primary external benefit of the protocol change is the reduction in energy consumption, which directly impacts a DLT's CO2 emissions. Using the QLR approach, which essentially implies doing a Chow-Break test for all the possible periods, I did not locate another relevant break in the sample. Given that a structural break was found at a 95% confidence level for Ethereum, hypotheses one and two must be re-examined. This will allow a more comprehensive description of the ecosystem's carbon footprint.

6.3.1 Reassessment of hypothesis 1

When accounting for the structural break, the same results for hypothesis 1, as not considering the break, were found. This affirms that the first hypothesis cannot be rejected, making the result more robust. It is possible to discern with certainty that, while the Ethereum blockchain was following a PoW approach, it presented a statistically different carbon emission per transaction compared to Bitcoin. The same was found for Ethereum following PoS, after the break. However, there was insufficient evidence to prove that the average change in emissions between DLTs differs in the log differences and considering the structural break. This implies that both PoW and PoS have a similar underlying trend represented by the log differences. A similar conclusion was drawn when not accounting for structural breaks.

6.3.2 Reassessment of hypothesis 2

The structural break also affects how the model derivation is applied. Because of the break, the data sample was split in two, one before and one after the event. The model underwent significant change before the break. The variables failed the stationarity test in absolute terms, but all variables were found stationary in their log differences. For Ethereum pre-break, the hashing power and price became significantly relevant in addition to energy consumption.

This follows the expectation from de Vries et al. (2022), who add more focus on externalities when accounting for a DLTs carbon footprint. This research's proposed model similarly suggests that a group running a PoW DLT must consider three variables and their polynomial properties to optimize carbon emission.

Ethereum, after the break, kept a similar model as when not considereding the break, in which only energy in its first polynomial was significant. This implies that there may be an omitted variable bias. Thus, more indirect variables could affect the PoS blockchain's carbon emissions per transaction trends.

The model for the ecosystem, with the NW standard errors, revealed that only Ethereum is relevant to explain the ecosystem's emissions. Thus, while Ethereum ran a PoW approach, energy hashing power and price were relevant. Once the soft fork happened, only energy continued to be relevant. As a result, the second hypothesis considering structural breaks, should be rejected.

The model in its VAR showed a worsening in its R-squared, in comparison to the NW. Thus, having a worse explanatory power. In the VAR case, all equations were found relevant. The second hypothesis continues to be rejected because, after the break, only energy is statistically relevant. The IRF found that shock in the ecosystem carbon footprint caused by itself have a permanent positive response. The Granger causality test found that there is a two-way Granger causality within the variables. When testing for IVs, no available variable was found relevant. Further, the unit root tests indicated that the residuals are non-stationary. Also, the VAR model equations cointegrate with each other. The VECM model shows an improvement in the R-squares. The cointegration equation from the VECM found that most of the time DLTs have a positive relationship with the ecosystem in the long turn. This implies that permanent impulses (e.g., regulation and forks), to have a lasting effect on the carbon footprint of the ecosystem, can target individual DLTs. This is because the long-term relationship caused by the cointegration will reach a new equilibrium that accounts for the exogenous sock.

6.4 Predictive power

The predictive power from the VECM derived during the second and third hypothesis was assessed using the DM test. The benchmark used for the DM test is a simple AR(1) model. The proposed VECM for Ethereum and the whole ecosystem did not have a statistically different average RMSE in comparison to the benchmarks. Implying that the predictive power of the proposed VECM does not add enough information to the relationship to outperform a model constructed with only the information of the lag of the exogenous variables. For Bitcoin's VECM, the DB highlights its weak predictive power. The VECM model has a higher and statistically different RMSE. Thus, despite the higher R-squares from the VECM, which suggest a superior explanatory power, they do not have a stronger explanatory power than the benchmark.

6.5 Answering the research question

All the hypotheses were derived to assist in answering the research question:

> *"What are the main drivers of the carbon footprint from the blockchain ecosystem and its externalities between 2019-2022?"*

According to this research, energy consumption is the primary driver of the carbon footprint for the public blockchain ecosystem. Energy consumption is the only variable relevant over the years 2019-2022. In most cases, energy consumption positively affects carbon emissions. Other variables were proven relevant to explain carbon emissions drivers in DLTs: price of native tokens and hashing power of DLTs. This research's models exhibit cointegration within the variables. Thus, permanent shocks in individual drivers will have a lasting effect on carbon emissions. However, the variables stopped being relevant because of the soft fork in Ethereum. Consequently, as it is known that other DLTs with a similar structure exist, those variables are still considered relevant for future DLT projects and regulations.

Based on the models given, it is possible to assume that as hashing power increases considerably, users with outdated technology (temporarily) stop mining to save costs. Consequently, after miners disactivated their node, the hashing power decreases. If the shift is significant enough, the DLT's carbon footprint decreases correspondingly. However, if the decline in hashing power does not pass a minimum threshold, the negative externality is that ecosystem's carbon footprint is predicted to increase. This is contrary to expectations.

Another key finding is that increases in prices of native tokens and the hashing power will decrease the carbon emissions in the long run due to an increase in the competitiveness in the environment. Increases in price drive up the overall hashing power for DLTs, simultaneously, increases in hashing power also decreases the number of users mining nodes. The result of this is the externality of miners purchasing more efficient equipment. This is a positive externality, while new equipment is expected to consume less energy and consequently, contributes to diminishing carbon footprints. Thus, if the price of native tokens and/or hashing power levels increase, and enough miners upgrade equipment, the ecosystem's carbon footprint is likely to decrease.

6.6 Limitations and recommendations

The primary limitation of this research stems from the scarcity of data available for small and private DLTs. In order to comprehend the intricacies associated with their carbon footprint, one must conduct thorough investigations that delve deep into the distinctions among different layers of blockchain technologies. Finding data without incurring excessive costs poses a significant challenge, although Bitcoin and Ethereum constitute approximately 80% of the public ecosystem (The White House, 2022).

It is crucial to acknowledge that the ecosystem modelling process, as how I developed it within this research, is an indirect approach to calculate its carbon emissions. An alternative study may find ways to employ a direct method for peer-reviewing the identified drivers. Therefore, future research should consider adopting multiple approaches to validate and enhance the understanding of carbon emissions in DLTs.

Furthermore, new research focusing on potential IVs is needed to address the endogeneity discovered while estimating the relationship between electricity consumption and carbon emissions in DLTs. Essentially, IVs act as proxy variables that are correlated with the endogenous variable of interest, but not directly correlated with the error term in the regression equation. One potential instrumental variable to explore in future research is the renewable energy mix from the geo-locations of miners. This variable could capture the exogenous variation in electricity sources and serve as a suitable instrument to address endogeneity concerns. By incorporating additional relevant IVs into the analysis, researchers can obtain more reliable and unbiased estimates of the causal relationship between electricity consumption and carbon emissions in DLTs, ultimately contributing to a more robust understanding of blockchain technologies' environmental impact.

Notably, energy consumption emerged as the sole pertinent variable throughout the entire data sample investigated. Therefore, the most effective strategy to reduce the carbon footprint of a DLT lies in adopting a PoS approach. According to this research' findings, a systematic and gradual transition from PoW to PoS consensus mechanisms would have positive externalities. I recommend regulatory entities to incentivize the aforementioned transition, for instance with tax cuts on utility costs. Ideally, stakeholders involved in DLTs should prioritize formulating and initiating regulatory actions, while involving experienced blockchain developers in the process. Control on energy consumption required per transaction is vital to foster blockchains.

This research presented variables that lost their relevance after the structural break, which should still be taken into account in future research or regulation, or DLT ventures. Data from the layers of Ethereum before the break can be used as a point of reference, because there are still similar DLTs existent or under development in the ecosystem.

Lastly, because of the competitive relationship between the variables in the long-run, future regulation should promote R&D in more efficient mining hardware and set policies to accelerate the adoption rate from miners towards more efficient hardware.

In conclusion, despite the aforementioned limitations, this research sheds light on the complexities associated with the carbon footprint of DLTs. It highlights the necessity for careful exploration of various blockchain technology layers and emphasizes the significance of energy consumption control.

REFERENCES

Alam, S., Shuaib, M., Khan, W. Z., Garg, S., Kaddoum, G., Hossain, M. S., & Zikria, Y. B. (2021). Blockchain-based initiatives: Current State and challenges. *Computer Networks*, *198*, [108395]. https://doi.org/10.1016/j.comnet.2021.108395

Alvi, S. T., Uddin, M. N., Islam, L., & Ahamed, S. (2022). DVTChain: A blockchain-based decentralized mechanism to ensure the security of Digital Voting System. *Journal of King Saud University - Computer and Information Sciences*, *34*(9), 6855–6871. https://doi.org/10.1016/j.jksuci.2022.06.014

Amarasinghe, S. L., Su, S., Dong, X., Zappia, L., Ritchie, M. E., & Gouil, Q. (2020). Opportunities and challenges in long-read sequencing data analysis. *Genome Biology*, *21*(1). https://doi.org/10.1186/s13059-020-1935-5

Best, R. de. (2023, May 9). *Bitcoin energy consumption 2023*. Statista. https://www.statista.com/statistics/881541/bitcoin-energy-consumption-transaction-comparison-visa/

Bruns, S. B., & Stern, D. I. (2018). Lag length selection and p-hacking in Granger causality testing: Prevalence and performance of meta-regression models. *Empirical Economics*, *56*(3), 797–830. https://doi.org/10.1007/s00181-018-1446-3

Cambridge Univeristy. (2021). *3rd global cryptoasset benchmarking study - CCAF publications*. Cambridge Judge Business School. https://www.jbs.cam.ac.uk/faculty-research/centres/alternative-finance/publications/3rd-global-cryptoasset-benchmarking-study/

Cambridge Univeristy. (2023). *Cambridge Bitcoin Electricity Consumption Index (CBECI)*. CCAF.io. https://ccaf.io/cbnsi/cbeci

Carbon Credits. (2022). *Tesla Carbon Credit Sales Jump by 116%*. Carbon Credits. https://carboncredits.com/tesla-regulatory-carbon-credit-sales-jumps-116/

Carbon Credits. (2023). *Tesla Carbon Credit Sales Reach record $1.78 billion in 2022*. Carbon Credits. https://carboncredits.com/tesla-carbon-credit-sales-reach-record-1-78-billion-in-2022/

Carter, N. (2021, May 6). *How much energy does Bitcoin actually consume?*. Harvard Business Review. https://hbr.org/2021/05/how-much-energy-does-bitcoin-actually-consume

Castillo, M. del. (2023, June 10). *Forbes Blockchain 50 2022*. Forbes. https://www.forbes.com/sites/michaeldelcastillo/2022/02/08/forbes-blockchain-50-2022/

CCRI. (2022). *CCRI - understanding your climate impact.* CCRI - Crypto Carbon Ratings Institute. https://carbon-ratings.com/

Cooling, S. (2021). *Proof-of-stake: Ethereum 2.0 ranks Worst Energy Consumption*. Yahoo! Finance. https://finance.yahoo.com/news/proof-stake-ethereum-2-0-153221744.html

Deloitte. (2018*). Crunch time IV blockchain for finance*. Deloitte. Retrieved January 9, 2022, from https://www2.deloitte.com/content/dam/Deloitte/global/Documents/Finance/gx-ft-crunch-time-blockchain-finance.pdf

de Vries, A., & Stoll, C. (2021). Bitcoin's growing e-waste problem. *Resources, Conservation and Recycling*, *175*, 105901. https://doi.org/10.1016/j.resconrec.2021.105901

de Vries, A., Gallersdörfer, U., Klaaßen, L., & Stoll, C. (2022). Revisiting bitcoin's carbon footprint. *Joule*, *6*(3), 498–502. https://doi.org/10.1016/j.joule.2022.02.005

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, *74*(366), 427. https://doi.org/10.2307/2286348

Digiconomist. (2023, February 2). Digiconomist Index. https://digiconomist.net/
Economic Times. (2022). *Bitcoin or gold? which is the future of investments?*. The Economic Times. https://economictimes.indiatimes.com/markets/cryptocurrency/bitcoin-or-gold-which-is-the-future-of-investments/articleshow/85858417.cms
European Commission. (2023). *EU Emissions Trading System (EU ETS)*. Climate Action. https://climate.ec.europa.eu/eu-action/eu-emissions-trading-system-eu-ets_en
Ethereum. (2023). *History and Forks of Ethereum*. Ethereum.org. https://ethereum.org/en/history/
Fitch. (2022). *Growing Crypto Mining Could Affect Energy Markets*. Fitch Ratings: Credit Ratings & Analysis for Financial Markets. https://www.fitchratings.com/research/corporate-finance/growing-crypto-mining-could-affect-energy-markets-21-02-2022
Frankenfield, J. (2022, October 10). *Hard fork: What it is in blockchain, how it works, why it happens*. Investopedia. https://www.investopedia.com/terms/h/hard-fork.asp/
Frankenfield, J. (2023, May 1). *Soft fork: What it is, how it works in cryptocurrency*. Investopedia. https://www.investopedia.com/terms/s/soft-fork.asp
Gallersdörfer, U., Klaaßen, L., & Stoll, C. (2020). Energy consumption of cryptocurrencies beyond bitcoin. *Joule*, *4*(9), 1843–1846. https://doi.org/10.1016/j.joule.2020.07.013
Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, *37*(3), 424. https://doi.org/10.2307/1912791
Ghosh, B., & Bouri, E. (2022). Is Bitcoin's carbon footprint persistent? Multifractal evidence and policy implications. *Entropy*, *24*(5), 647. https://doi.org/10.3390/e24050647
Gulli, A. (2020). (Un)Sustainability of Bitcoin Mining. Rutgers Computer and Technology Law Journal, 46(1), 95–123.
HM Treasury. (n.d.). *Statutory framework for the Early Years foundation stage - gov.uk*. Gov UK. https://assets.publishing.service.gov.uk/government/
Hook, L. (2020, September 22). *China pledges to be "carbon-neutral" by 2060*. Financial Times. https://www.ft.com/content/730e4f7d-3df0-45e4-91a5-db4b3571f353
IBM. (2020). *Advancing global trade with blockchain advancing*. IBM Institute for Business Value. https://www.ibm.com/downloads/cas/WVDE0MXG
Klaaßen, L., Gallersdörfer, U., & Stoll, C. (2020). *Energy consumption of cryptocurrencies beyond bitcoin*. Joule. https://pubmed.ncbi.nlm.nih.gov/32838201/
Kwiatkowski, D.; Phillips, P. C. B.; Schmidt, P.; Shin, Y. (1992). "Testing the null hypothesis of stationarity against the alternative of a unit root". Journal of Econometrics. 54 (1–3): 159–178. https://doi:10.1016/0304-4076(92)90104-Y
Lashkari, B., & Musilek, P. (2021). A comprehensive review of Blockchain Consensus Mechanisms. *IEEE Access*, *9*, 43620–43652. https://doi.org/10.1109/access.2021.3065880
Li, J., Li, N., Peng, J., Cui, H., & Wu, Z. (2019). Energy consumption of cryptocurrency mining: A study of electricity consumption in mining cryptocurrencies. *Energy*, *168*, 160–168. https://doi.org/10.1016/j.energy.2018.11.046
Loreen, M., Mangle, A., & Wimmer, H. (2021). Awareness of blockchain usage, structure, & generation of platforms energy consumption: Working towards a greener blockchain. *Issues In Information Systems*. https://doi.org/10.48009/1_iis_2021_114-123
McDonald, K. (2022). *Ethereum emissions: A bottom-up estimate*. Arxiv. https://arxiv.org/pdf/2112.01238v3.pdf
Merriam-Webster. (2023). *Ledger definition & meaning*. Merriam-Webster. https://www.merriam-webster.com/dictionary/ledger
Nakamoto, S. (2008). *Bitcoin: A peer-to-peer electronic cash system*. bitcoin.org. Retrieved January 9, 2022, from http://satoshinakamoto.me/bitcoin.pdf

NYDIG. (2021). *Bitcoin Net Zero: Nydig - Bitcoin for all*. NYDIG. https://nydig.com/bitcoin-net-zero

Oxford. (2022). *Oxford Blockchain Strategy Programme | Saïd Business School*. Oxford. https://www.sbs.ox.ac.uk/programmes/executive-education/online-programmes/oxford-blockchain-strategy-programme

Pennella, L. (2023). *Proof-of-stake (POS)*. ethereum.org. https://ethereum.org/en/developers/docs/consensus-mechanisms/pos/

Phillips, P. C. B.; Perron, P. (1988). "Testing for a Unit Root in Time Series Regression" (PDF). Biometrika. 75 (2): 335–346. https://doi:10.1093/biomet/75.2.335

Platt, M., Sedlmeir, J., Platt, D., Xu, J., Tasca, P., Vadgama, N., & Ibanez, J. I. (2021). The energy footprint of blockchain consensus mechanisms beyond proof-of-work. *2021 IEEE 21st International Conference on Software Quality, Reliability and Security Companion (QRS-C)*. https://doi.org/10.1109/qrs-c55045.2021.00168

Quiroz-Gutierrez, M. (2022, January 5). *Crypto is fully banned in China and 8 other countries*. Fortune. https://fortune.com/2022/01/04/crypto-banned-china-other-countries/

Sarkar, A. (2022). *The merge brings down Ethereum's network power consumption by over 99.9%*. Cointelegraph. https://cointelegraph.com/news/the-merge-brings-down-ethereum-s-network-power-consumption-by-over-99-9

Sarkodie, S. A., Ahmed, M. Y., & Leirvik, T. (2022). Trade volume affects bitcoin energy consumption and carbon footprint. *Finance Research Letters*, *48*, 102977. https://doi.org/10.1016/j.frl.2022.102977

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*(2). https://doi.org/10.1214/aos/1176344136

Sergeenkov, A. (2023, May 11). *China crypto bans: A complete history*. CoinDesk Latest Headlines RSS. https://www.coindesk.com/learn/china-crypto-bans-a-complete-history/

Sims, C. A. (1980). Macroeconomics and reality. *Econometrica*, *48*(1), 1. https://doi.org/10.2307/1912017

Singh, O. (2022, October 5). *What is bitcoin hash rate and why does it matter?*. Cointelegraph. https://cointelegraph.com/explained/what-is-bitcoin-hash-rate-and-why-does-it-matter

Sipthorpe, A., Brink, S., Van Leeuwen, T., & Staffell, I. (2022). Blockchain solutions for carbon markets are nearing maturity. *One Earth*, *5*(7), 779–791. https://doi.org/10.1016/j.oneear.2022.06.004

The Guardian. (2022, July 16). *Energy use from US cryptomining firms is contributing to rising utility bills*. The Guardian. https://www.theguardian.com/technology/2022/jul/16/cryptomining-energy-use-investigation-carbon-emissions

The White House. (2022). *FACT SHEET: Climate and Energy Implications of Crypto-Assets in the United States*. Retrieved from https://www.whitehouse.gov/ostp/news-updates/2022/09/08/fact-sheet-climate-and-energy-implications-of-crypto-assets-in-the-united-states/

UN. (2016). *List of parties that signed the Paris Agreement on 22 April*. United Nations. https://www.un.org/sustainabledevelopment/blog/2016/04/parisagreementsingatures/

UN. (2023). *EU taxonomy for Sustainable Activities*. UN - Finance. https://finance.ec.europa.eu/sustainable-finance/tools-and-standards/eu-taxonomy-sustainable-activities_en

Vieira, L. I., & Laurini, M. P. (2022, December 23). *Time-varying higher moments in Bitcoin - Digital Finance*. SpringerLink. https://link.springer.com/article/10.1007/s42521-022-00072-8

Wang, K. (2021). (rep.). *Measuring Global Crypto Users, A Study to Measure Market Size Using On-Chain Metrics* (pp. 1–15). New York, NY: Crypto.com.

Yang, R., Wakefield, R., Lyu, S., Jayasuriya, S., Han, F., Yi, X., Yang, X., Amarasinghe, G., & Chen, S. (2020). Public and private blockchain in construction business process and information integration. *Automation in Construction*, *118*, 103276. https://doi.org/10.1016/j.autcon.2020.103276

Zapata, H. O., & Rambaldi, A. N. (1997). Monte Carlo Evidence on cointegration and causation. *Oxford Bulletin of Economics and Statistics*, *59*(2), 285–298. https://doi.org/10.1111/1468-0084.00065

Zerofy. (2022). *Typical household CO2 emissions and their breakdown into spending categories*. Zerofy. https://www.zerofy.net/2022/04/04/household-co2-emissions.html

APPENDIX A TABLES

List of tables:

.

Table 1A *Ethereum fork history*

Fork name	Date	Description
Muir Glacier	02/01/2020	decrease waiting times in sending transactions and using apps
staking deposit contract deployed	14/10/2020	implemented staking in the ecosystem
Beacon Chin Genesis	01/12/2020	skating chain started
Berlin	15/04/2021	optimization of transaction costs
London	05/08/2021	further updated on transaction costs
Altair	27/10/2021	increased monitoring and penalties on validators
Arrow Glacier	09/12/2021	decrease waiting times in sending transactions and using apps
Gray Glacier	30/06/2022	decrease waiting times in sending transactions and using apps
Bellatrix	06/09/2022	preparing for the merge

Note. Forks of 2023 are not considered given the scope of this research.

Table 2A *Results Dickey-Fuller test for unit root on cleaned variables*

Null hypothesis	P-value
Energy BTC Random Walk Without Drift	0.000
Price BTC Random Walk Without Drift	0.671
Hash BTC Random Walk Without Drift	0.000
CO2 BTC Random Walk Without Drift	0.000
Energy ETH Random Walk Without Drift	0.307
E-waste Random Walk Without Drift	0.847
Price ETH Random Walk Without Drift	0.542
Hash ETH Random Walk Without Drift	0.762
CO2 ETH Random Walk Without Drift	0.141
CO2 Ecosystem Random Walk Without Drift	0.141
Energy ETH Random Walk Without Drift Before Break	0.691
E-waste Random Walk Without Drift Before Break	0.692
Price ETH Random Walk Without Drift Before Break	0.581
Hashing power ETH Random Walk Without Drift Before Break	0.940
CO2 ETH Random Walk Without Drift Before Break	0.120
Energy ETH Random Walk Without Drift After Break	0.060
E-waste Random Walk Without Drift After Break	0.530
Price ETH Random Walk Without Drift After Break	0.243
Hashing power ETH Random Walk Without Drift After Break	0.056
CO2 ETH Random Walk Without Drift After Break	0.173

Note. BTC stands for Bitcoin, and ETH for Ethereum. CO2 relates to carbon emission. The log difference represents the change in the CO2 levels.

Table 3A *Results Dickey-Fuller test for unit root on log differences*

Null hypothesis	P-value
Energy BTC Random Walk Without Drift	0.000
Price BTC Random Walk Without Drift	0.000
Hashing power BTC Random Walk Without Drift	0.000
CO2 BTC Random Walk Without Drift	0.000
Energy ETH Random Walk Without Drift	0.000
E-waste Random Walk Without Drift	0.000
Price ETH Random Walk Without Drift	0.000

Hashing power ETH Random Walk Without Drift	0.000
CO2 ETH Random Walk Without Drift	0.000
CO2 Ecosystem Random Walk Without Drift	0.000
Energy ETH Random Walk Without Drift Before Break	0.000
E-waste Random Walk Without Drift Before Break	0.000
Price ETH Random Walk Without Drift Before Break	0.000
Hashing power ETH Random Walk Without Drift Before Break	0.000
CO2 ETH Random Walk Without Drift Before Break	0.000
Energy ETH Random Walk Without Drift After Break	0.000
E-waste Random Walk Without Drift After Break	0.000
Price ETH Random Walk Without Drift After Break	0.000
Hashing power ETH Random Walk Without Drift After Break	0.000
CO2 ETH Random Walk Without Drift After Break	0.000

Note. *Note*. BTC stands for Bitcoin, and ETH for Ethereum. CO2 relates to carbon emission. The log difference represents the change in the CO2 levels.

Table 4A *Results Phillips-Perron test for unit root on cleaned variables*

Null hypothesis	P-value
Energy BTC Random Walk Without Drift	0.045
Price BTC Random Walk Without Drift	0.665
Hashing power BTC Random Walk Without Drift	0.000
CO2 BTC Random Walk Without Drift	0.000
Energy ETH Random Walk Without Drift	0.860
E-waste Random Walk Without Drift	0.322
Price ETH Random Walk Without Drift	0.607
Hashing power ETH Random Walk Without Drift	0.627
CO2 ETH Random Walk Without Drift	0.918
CO2 Ecosystem Random Walk Without Drift	0.743
Energy ETH Random Walk Without Drift Before Break	0.691
E-waste Random Walk Without Drift Before Break	0.895
Price ETH Random Walk Without Drift Before Break	0.590
Hashing power ETH Random Walk Without Drift Before Break	0.940
CO2 ETH Random Walk Without Drift Before Break	0.682
Energy ETH Random Walk Without Drift After Break	0.006
E-waste Random Walk Without Drift After Break	0.196
Price ETH Random Walk Without Drift After Break	0.243

Hashing power ETH Random Walk Without Drift After Break	0.008
CO2 ETH Random Walk Without Drift After Break	0.608

Note. BTC stands for Bitcoin, and ETH for Ethereum. CO2 relates to carbon emission. The log difference represents the change in the CO2 levels.

Table 5A *Results Phillips-Perron test for unit root on log differences*

Null hypothesis	P-value
Energy BTC Random Walk Without Drift	0.000
Price BTC Random Walk Without Drift	0.000
Hashing power BTC Random Walk Without Drift	0.000
CO2 BTC Random Walk Without Drift	0.000
Energy ETH Random Walk Without Drift	0.000
E-waste Random Walk Without Drift	0.000
Price ETH Random Walk Without Drift	0.000
Hashing power ETH Random Walk Without Drift	0.000
CO2 ETH Random Walk Without Drift	0.000
CO2 Ecosystem Random Walk Without Drift	0.000
Energy ETH Random Walk Without Drift Before Break	0.000
E-waste Random Walk Without Drift Before Break	0.000
Price ETH Random Walk Without Drift Before Break	0.000
Hashing power ETH Random Walk Without Drift Before Break	0.000
CO2 ETH Random Walk Without Drift Before Break	0.000
Energy ETH Random Walk Without Drift After Break	0.000
E-waste Random Walk Without Drift After Break	0.000
Price ETH Random Walk Without Drift After Break	0.000
Hashing power ETH Random Walk Without Drift After Break	0.000
CO2 ETH Random Walk Without Drift After Break	0.000

Note. BTC stands for Bitcoin, and ETH for Ethereum. CO2 relates to carbon emission. The log difference represents the change in the CO2 levels.

Table 6A *Results Kwiatkowski-Phillips-Schmidt-Shin test for unit root on cleaned variables*

Null hypothesis	Test statistic
Energy BTC Random Walk Without Drift	9.67 (0 lag)
Price BTC Random Walk Without Drift	15.2. (0 lag)
Hashing power BTC Random Walk Without Drift	1.101 (0 lag)
CO2 BTC Random Walk Without Drift	5.2 (0 lag)

Energy ETH Random Walk Without Drift	9.16 (0 lag)
E-waste Random Walk Without Drift	11.70 (0 lag)
Price ETH Random Walk Without Drift	12.70 (0 lag)
Hashing power ETH Random Walk Without Drift	7.88 (0 lag)
CO2 ETH Random Walk Without Drift	7.68 (0 lag)
CO2 Ecosystem Random Walk Without Drift	2.783 (0 lag)
Energy ETH Random Walk Without Drift Before Break	24.7 (0 lag)
E-waste Random Walk Without Drift Before Break	8.22(0 lag)
Price ETH Random Walk Without Drift Before Break	10.1 (0 lag)
Hashing power ETH Random Walk Without Drift Before Break	21.7 (0 lag)
CO2 ETH Random Walk Without Drift Before Break	23.7 (0 lag)
Energy ETH Random Walk Without Drift After Break	0.611 (0 lag)
E-waste Random Walk Without Drift After Break	0.410 (0 lag)
Price ETH Random Walk Without Drift After Break	0.479 (0 lag)
Hashing power ETH Random Walk Without Drift After Break	0.751 (0 lag)
CO2 ETH Random Walk Without Drift After Break	0.758 (0 lag)

Note. BTC stands for Bitcoin, and ETH for Ethereum. CO2 relates to carbon emission. The log difference represents the change in the CO2 levels. The critical value for the KPSS test is 0.146

Table 7A *Results Kwiatkowski-Phillips-Schmidt-Shin test for unit root on log differences*

Null hypothesis	Test statistic
Energy BTC Random Walk Without Drift	0.076 (18 lag)
Price BTC Random Walk Without Drift	0.143 (22 lag)
Hashing power BTC Random Walk Without Drift	0.025 (22 lag)
CO2 BTC Random Walk Without Drift	0.037 (22 lag)
Energy ETH Random Walk Without Drift	0.149 (0 lag)
E-waste Random Walk Without Drift	0.07 (0 lag)
Price ETH Random Walk Without Drift	0.167 (2 lag)
Hashing power ETH Random Walk Without Drift	0.185 (0 lag)
CO2 ETH Random Walk Without Drift	0.084(0 lag)
CO2 Ecosystem Random Walk Without Drift	0.107 (0 lag)
Energy ETH Random Walk Without Drift Before Break	0.063 (22 lag)
E-waste Random Walk Without Drift Before Break	0.072 (0 lag)
Price ETH Random Walk Without Drift Before Break	0.167 (2 lag)
Hashing power ETH Random Walk Without Drift Before Break	0.266 (15 lag)
CO2 ETH Random Walk Without Drift Before Break	0.025 (0 lag)
Energy ETH Random Walk Without Drift After Break	0.135 (0 lag)

E-waste Random Walk Without Drift After Break	0.273 (0 lag)
CO2 Ecosystem Random Walk Without Drift After Break	0.018 (11 lag)
CO2 Ecosystem Random Walk Without Drift Before Break	0.065 (19 lag)
Price ETH Random Walk Without Drift After Break	0.062 (11 lag)
Hashing power ETH Random Walk Without Drift After Break	0.113 (11 lag)
CO2 ETH Random Walk Without Drift After Break	0.074 (0 lag)

Note. BTC stands for Bitcoin, and ETH for Ethereum. CO2 relates to carbon emission. The log difference represents the change in the CO2 levels. The critical value for the KPSS test is 0.146

Table 8A *BTC Energy*

Variable	(1) OLS	(2) NW	(3) OLS	(4) NW	(5) OLS	(6) OLS	(7) NW
Energy BTC	0.971^{***}	0.971^{***}	0.968^{***}	0.968^{***}	0.978^{***}	0.969^{***}	0.969^{***}
	(0.006)	(0.007)	(0.006)	(0.007)	(0.008)	(0.006)	(0.007)
Energy BTC 2nd Polynomial			0.077^{*}	0.077			
			(0.031)	(0.043)			
Energy BTC 3rd Polynomial					-0.119		
					(0.100)		
Energy BTC 4th Polynomial						0.461^{**}	0.461^{**}
						(0.163)	(0.164)
Constant	0.000	0.000	-0.001	-0.001	0.000	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	1289	1289	1289	1289	1289	1289	1289
R^2	0.950		0.950		0.950	0.950	

Note. Standard errors in parentheses. $^{*}p < 0.05$. $^{**}p < 0.01$. $^{***}p < 0.001$.

Table 9A *Hashing Power BTC*

Variable	(1) OLS	(2) OLS	(3) OLS	(4) OLS
Energy BTC	0.969^{***}	0.969^{***}	0.969^{***}	0.969^{***}
	(0.006)	(0.006)	(0.006)	(0.006)
Energy BTC 4rd Polynomial	0.461^{**}	0.454^{**}	0.461^{**}	0.460^{**}
	(0.163)	(0.163)	(0.163)	(0.163)
Hashing power BTC	0.000			
	(0.000)			
Hashing power BTC 2nd Polynomial		-0.000		
		(0.000)		
Hashing power BTC 3rd Polynomial			0.000	
			(0.000)	
Hashing power BTC 4th Polynomial				0.000
				(0.000)
Constant	0.000	0.000	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)
Observations	1289	1289	1289	1289
R^2	0.950	0.950	0.950	0.950

Note. Standard errors in parentheses. $^{*}p < 0.05$. $^{**}p < 0.01$. $^{***}p < 0.001$.

Table 10A *E-waste BTC*

Variable	(1) OLS	(2) NW	(3) OLS	(4) OLS	(5) OLS
Energy BTC	0.969^{***}	0.969^{***}	0.971^{***}	0.971^{***}	0.971^{***}

	(0.006)	(0.007)	(0.006)	(0.006)	(0.006)
Energy BTC 4	0.467**	0.467**			
	(0.163)	(0.163)			
E-waste	0.079*	0.079			
	(0.040)	(0.049)			
E-waste 2nd Polynomial			1.353 (0.868)		
E-waste 3rd Polynomial				7.251 (10.110)	
E-waste 4th Polynomial					59.330 (102.500)
Constant	0.000	0.000	0.000	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	1289	1289	1289	1289	1289
R^2	0.950		0.950	0.950	0.950

Note. Standard errors in parentheses. $^{*}p < 0.05$. $^{**}p < 0.01$. $^{***}p < 0.001$.

Table 11A *Price BTC*

Variable	(1) OLS	(2) OLS	(3) OLS	(4) OLS
Energy BTC	0.970***	0.969***	0.969***	0.969***
	(0.006)	(0.006)	(0.006)	(0.006)
Energy BTC 4th Polynomial	0.462** (0.163)	0.462** (0.163)	0.462** (0.163)	0.462** (0.163)
Price BTC	0.003			
	(0.002)			
Price BTC 2nd Polynomial		0.000 (0.000)		
Price BTC 3rd Polynomial			0.000 (0.000)	
Price BTC 4th Polynomial				0.000 (0.000)
Constant	0.000	0.000	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)
Observations	1289	1289	1289	1289
R^2	0.950	0.950	0.950	0.950

Note. Standard errors in parentheses. $^{*}p < 0.05$. $^{**}p < 0.01$. $^{***}p < 0.001$.

Table 12A *Energy ETH*

Variable	(1) OLS	(2) OLS	(3) OLS	(4) NW	(5) NW	(6) NW	(7) OLS
Energy ETH	0.371***	0.845***	1.047***	0.371***	0.845***	1.047***	1.062***
	(0.110)	(0.248)	(0.227)	(0.088)	(0.183)	(0.199)	(0.201)
Energy ETH 2nd Polynomial		-0.116** (0.040)	-0.549*** (0.114)		-0.116** (0.039)	-0.549** (0.173)	-0.099 (0.719)
Energy ETH 3rd Polynomial			0.072*** (0.013)			0.073* (0.028)	-0.159 (0.359)
Energy ETH 4th Polynomial							0.027 (0.041)

Constant	0.016	0.018	0.022	0.016	0.018	0.022	0.019
	(0.017)	(0.018)	(0.019)	(0.018)	(0.018)	(0.018)	(0.019)
Observations	1289	1289	1289	1289	1289	1289	1289
R^2				0.014	0.020	0.025	0.026

Note. Standard errors in parentheses. $^{*}p < 0.05$. $^{**}p < 0.01$. $^{***}p < 0.001$.

Table 13A *Hashing Power ETH*

Variable	(1) OLS	(2) OLS	(3) OLS	(4) OLS
Energy ETH	1.042^{***}	1.051^{***}	1.055^{***}	1.054^{***}
	(0.199)	(0.200)	(0.200)	(0.200)
Energy ETH 2nd Polynomial	-0.507^{**}	-0.333	-0.121	-0.129
	(0.177)	(0.677)	(0.715)	(0.715)
Energy ETH 3rd Polynomial	0.058	0.002	-0.067	-0.064
	(0.031)	(0.216)	(0.228)	(0.228)
Hashing power ETH	0.975			
	(0.862)			
Hashing power ETH 2nd Polynomial		3.005		
		(9.114)		
Hashing power ETH 3rd Polynomial			4.212	
			(6.832)	
Hashing power ETH 4th Polynomial				2.939
				(4.862)
Constant	0.023	0.019	0.019	0.019
	(0.018)	(0.020)	(0.019)	(0.019)
Observations	1289	1289	1289	1289
R^2	0.026	0.025	0.026	0.026

Note. Standard errors in parentheses. $^{*}p < 0.05$. $^{**}p < 0.01$. $^{***}p < 0.001$.

Table 14A *Dummy Variable Protocol ETH*

Variable	(1) OLS
Energy ETH	1.059^{***}
	(0.198)
Energy ETH 2nd Polynomial	-0.565^{**}
	(0.172)
Energy ETH 3rd Polynomial	0.075^{**}
	(0.028)
Protocol	0.273
	(0.072)
Constant	0.005
	(0.018)
Observations	1289
R^2	0.036

Note. Standard errors in parentheses. $^{*}p < 0.05$. $^{**}p < 0.01$. $^{***}p < 0.001$.

Table 15A *E-waste ETH*

Variable	(1) OLS	(2) OLS	(3) OLS	(4) OLS
Energy ETH	1.046^{***}	1.047^{***}	1.047^{***}	1.047^{***}
	(0.199)	(0.199)	(0.199)	(0.199)

Energy ETH 2nd Polynomial	-0.548**	-0.548**	-0.549**	-0.549**
	(0.173)	(0.173)	(0.173)	(0.173)
Energy ETH 3rd Polynomial	0.072*	0.072*	0.072*	0.072*
	(0.028)	(0.028)	(0.028)	(0.028)
E-waste	0.106			
	(0.909)			
E-waste 2nd Polynomial		-8.691		
		(19.88)		
E-waste 3rd Polynomial			-9.181	
			(231.1)	
E-waste 4th Polynomial				-361.8
				(2344.5)
Constant	0.022	0.026	0.023	0.023
	(0.018)	(0.019)	(0.018)	(0.018)
Observations	1289	1289	1289	1289
R^2	0.025	0.025	0.025	0.025

Note. Standard errors in parentheses. $^{*}p < 0.05$. $^{**}p < 0.01$. $^{***}p < 0.001$.

Table 16A *Price ETH*

Variable	(1) OLS	(2) OLS	(3) OLS	(4) OLS
Energy ETH	1.072***	1.061***	1.069***	1.058***
	(0.204)	(0.201)	(0.202)	(0.201)
Energy ETH 2nd	-0.561**	-0.553**	-0.555**	-0.551**
Polynomial	(0.174)	(0.173)	(0.173)	(0.173)
Energy ETH 3rd	0.074**	0.073**	0.073**	0.072*
Polynomial	(0.028)	(0.028)	(0.028)	(0.028)
Price ETH	-0.202			
	(0.352)			
Price ETH 2nd Polynomial		-1.197		
		(2.433)		
Price ETH 3rd Polynomial			-5.809	
			(8.661)	
Price ETH 4th Polynomial				-11.190
				(26.14)
Constant	0.022	0.026	0.023	0.023
	(0.018)	(0.019)	(0.018)	(0.018)
Observations	1289	1289	1289	1289
R^2	0.026	0.026	0.026	0.025

Note. Standard errors in parentheses. $^{*}p < 0.05$. $^{**}p < 0.01$. $^{***}p < 0.001$.

Table 17A *Energy ETH Before the Break*

Variable	(1) OLS	(2) NW	(3) OLS	(4) NW	(5) OLS	(6) NW	(7) OLS	(8) NW
Energy ETH	0.846***	0.846***	0.854***	0.854***	0.896***	0.896***	0.856***	0.856***
	(0.013)	(0.023)	(0.013)	(0.017)	(0.017)	(0.028)	(0.013)	(0.017)
Energy ETH 2nd Polynomial			-0.225***	-0.225				
			(0.054)	(0.134)				
Energy ETH 3rd Polynomial					-0.789***	-0.789		
					(0.177)	(0.522)		
Energy ETH 4th Polynomial							-1.827***	-1.827*
							(0.303)	(0.740)
Constant	0.001	0.001	0.002*	0.002*	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	1198	1198	1198	1198	1198	1198	1198	1198
R2	0.790		0.793		0.793		0.796	

Note. Standard errors in parentheses. $^{*}p < 0.05$. $^{**}p < 0.01$. $^{***}p < 0.001$.

Table 18A *Hashing Power ETH Before the Break*

Variable	(1) OLS	(2) NW	(3) OLS	(4) NW	(5) OLS	(6) OLS
Energy ETH	0.851***	0.851***	0.851***	0.851***	0.851***	0.851***
	(0.011)	(0.017)	(0.011)	(0.017)	(0.011)	(0.011)
Energy ETH 4th Polynomial	-1.932***	-1.932*	-1.934***	-1.934*	-1.934***	-1.934***
	(0.257)	(0.807)	(0.257)	(0.804)	(0.257)	(0.257)
Hashing power ETH	0.915***	0.915***	0.914***	0.914***	0.929***	0.915***
	(0.042)	(0.046)	(0.043)	(0.046)	(0.064)	(0.043)
Hashing power ETH 2nd Polynomial			-2.718*	-2.718*	-2.654	-2.424
			(1.361)	(1.345)	(1.378)	(2.567)
Hashing power ETH 3rd Polynomial					-10.040	
					(33.530)	
Hashing power ETH 4th Polynomial						-94.470
						(700.500)
Constant	0.003**	0.003**	0.004***	0.004***	0.004***	0.004**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	1198	1198	1198	1198	1198	1198
R^2	0.853		0.854		0.854	0.854

Note. Standard errors in parentheses. $^{*}p < 0.05$. $^{**}p < 0.01$. $^{***}p < 0.001$.

Table 19A *E-waste ETH Before the Break*

Variable	(1) OLS	(2) OLS	(3) OLS	(4) OLS
Energy ETH	0.850***	0.851***	0.851***	0.851***
	(0.011)	(0.011)	(0.011)	(0.011)
Energy ETH 4th Polynomial	-1.912***	-1.931***	-1.932***	-1.933***
	(0.257)	(0.257)	(0.257)	(0.257)
Hashing power ETH	0.914***	0.914***	0.914***	0.914***
	(0.042)	(0.042)	(0.042)	(0.042)
Hashing power ETH 2nd Polynomial	-2.643	-2.711*	-2.716*	-2.712*
	(1.362)	(1.362)	(1.362)	(1.361)
E-waste	0.052			

	(0.045)			
E-waste 2nd Polynomial		-0.181		
		(0.983)		
E-waste 3rd Polynomial			1.428	
			(11.340)	
E-waste 4th Polynomial				-52.140
				(114.800)
Constant	0.004***	0.004***	0.004***	0.004***
	(0.001)	(0.001)	(0.001)	(0.001)
Observations	1198	1198	1198	1198
R^2	0.854	0.854	0.854	0.854

Note. Standard errors in parentheses. $^{*}p < 0.05$. $^{**}p < 0.01$. $^{***}p < 0.001$.

Table 20A *Price ETH Before the Break*

Variable	(1) OLS	(2) NW	(3) OLS	(4) OLS	(5) OLS
Energy ETH	0.897***	0.897***	0.896***	0.897***	0.898***
	(0.009)	(0.011)	(0.009)	(0.009)	(0.009)
Energy ETH 4th Polynomial	-1.220***	-1.220*	-1.373***	-1.175***	-1.067***
	(0.214)	(0.493)	(0.234)	(0.256)	(0.265)
Hashing power ETH	0.959***	0.959***	0.958***	0.959***	0.960***
	(0.035)	(0.038)	(0.035)	(0.035)	(0.035)
Hashing power ETH 2nd Polynomial	-2.317*	-2.317*	-2.326*	-2.316*	-2.323*
	(1.121)	(1.145)	(1.120)	(1.121)	(1.121)
Price ETH	-0.353***	-0.353***	-0.357***	-0.350***	-0.351***
	(0.015)	(0.035)	(0.015)	(0.018)	(0.015)
Price ETH 2nd Polynomial			0.177		
			(0.112)		
Price ETH 3rd Polynomial				-0.163	
				(0.508)	
Price ETH 4th Polynomial					-1.324
					(1.359)
Constant	0.003**	0.003**	0.003**	0.003**	0.003**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	1198	1198	1198	1198	1198
R^2	0.901		0.901	0.901	0.901

Note. Standard errors in parentheses. $^{*}p < 0.05$. $^{**}p < 0.01$. $^{***}p < 0.001$.

Table 21A *Energy ETH After the Break*

Variable	(1) OLS	(2) NW	(3) OLS	(4) OLS	(5) OLS
Energy ETH	0.233	0.233**	0.782	0.409	0.314
	(0.367)	(0.069)	(1.307)	(0.967)	(0.835)
Energy ETH 2nd Polynomial			-0.112		
			(0.256)		
Energy ETH 3rd Polynomial				-0.007	
				(0.034)	
Energy ETH 4th Polynomial					-0.001
					(0.005)
Constant	0.241	0.241	0.242	0.240	0.240
	(0.256)	(0.253)	(0.257)	(0.258)	(0.258)
Observations	90	90	90	90	90
R^2	0.005		0.007	0.005	0.005

Note. Standard errors in parentheses.
$^{*}p < 0.05$. $^{**}p < 0.01$. $^{***}p < 0.001$.

Table 22A *Hashing Power ETH After the Break*

Variable	(1) OLS	(2) OLS	(3) OLS	(4) OLS
Energy ETH	0.267 (0.755)	0.249 (0.716)	0.248 (0.715)	0.248 (0.715)
Hashing power ETH	-0.183 (3.537)			
Hashing power ETH 2nd Polynomial		-0.062 (2.389)		
Hashing power ETH 3rd Polynomial			-0.042 (1.697)	
Hashing power ETH 4th Polynomial				-0.030 (1.207)
Constant	0.240 (0.258)	0.240 (0.258)	0.240 (0.258)	0.240 (0.258)
Observations	90	90	90	90
R^2	0.005	0.005	0.005	0.005

Note. Standard errors in parentheses. $^{*}p < 0.05$. $^{**}p < 0.01$. $^{***}p < 0.001$.

Table 23A *E-waste ETH After the Break*

Variable	(1) OLS	(2) OLS	(3) OLS	(4) OLS
Energy ETH	0.242 (0.373)	0.233 (0.369)	0.232 (0.369)	0.232 (0.369)
E-waste	4.242 (22.800)			
E-waste 2nd Polynomial		-216.300 (696.300)		
E-waste 3rd Polynomial			852.000 (12275.400)	
E-waste 4th Polynomial				-16384.100 (213403.300)
Constant	0.244 (0.258)	0.268 (0.273)	0.243 (0.259)	0.243 (0.260)
Observations	90	90	90	90
R^2	0.005	0.006	0.005	0.005

Note. Standard errors in parentheses. $^{*}p < 0.05$. $^{**}p < 0.01$. $^{***}p < 0.001$.

Table 24A *Price ETH After the Break*

Variable	(1) OLS	(2) OLS	(3) OLS	(4) OLS
Energy ETH	0.192 (0.376)	0.248 (0.374)	0.236 (0.372)	0.237 (0.370)
Price ETH	2.963 (5.454)			
Price ETH 2nd Polynomial		-11.640 (44.510)		
Price ETH 3rd Polynomial			-18.980 (250.300)	
Price ETH 4th Polynomial				-290.500 (1436.000)

Constant	0.238 (0.257)	0.266 (0.275)	0.242 (0.259)	0.251 (0.263)
Observations	90	90	90	90
R^2	0.008	0.005	0.005	0.005

Note. Standard errors in parentheses. $^{*}p < 0.05$. $^{**}p < 0.01$. $^{***}p < 0.001$.

Table 25A Results Breusch-Pagan/Cook-Weisberg test for heteroskedascity

Equation	P-value
ETH CO2, ETH Energy	0
ETH CO2, ETH Energy, ETH Energy2	0
ETH CO2, ETH Energy, ETH Energy2, ETH Energy3	0
BTC CO2, BTC Energy	0
BTC CO2, BTC Energy, BTC Energy2	0
BTC CO2, BTC Energy, BTC Energy4	0
BTC CO2, BTC Energy, BTC Energy4, E-waste	0
Ecosystem CO2, BTC CO2, ETH CO2	0
ETH CO2, ETH Energy before break	0
ETH CO2, ETH Energy, ETH Energy2 before the break	0
ETH CO2, ETH Energy, ETH Energy3 before the break	0
ETH CO2, ETH Energy, ETH Energy4 before the break	0
ETH CO2, ETH Energy, ETH Energy4, ETH Hashing1 before the break	0
ETH CO2, ETH Energy, ETH Energy4, ETH Hashing1, ETH Hashing2 before the break	0
ETH CO2, ETH Energy, ETH Energy4, ETH Hashing1, ETH Hashing2, ETH Price before the break	0
ETH CO2, ETH Energy after the break	0

Note. BTC stands for Bitcoin, and ETH for Ethereum. CO2 relates to carbon emission. The log difference represents the change in the CO2 levels. The number represents the polynomials. Before and after break refers to the structural break.

Table 26A Results Breusch-Pagan/Cook-Weisberg Lagrange Multiplier test for AC

Equation	P-value
ETH CO2, Eth energy	0
ETH CO2, ETH Energy, ETH Energy2	0
ETH CO2, ETH Energy, ETH Energy2, ETH Energy3	0
BTC CO2, BTC Energy	0
BTC CO2, BTC Energy, BTC Energy2	0
BTC CO2, BTC Energy, BTC Energy4	0
BTC CO2, BTC Energy, BTC Energy4, E-waste	0
Ecosystem CO2, BTC CO2, ETH CO2	0
ETH CO2, ETH Energy before break	0
ETH CO2, ETH Energy, ETH Energy2 before the break	0
ETH CO2, ETH Energy, ETH Energy3 before the break	0
ETH CO2, ETH Energy, ETH Energy4 before the break	0
ETH CO2, ETH Energy, ETH Energy4, ETH Hashing1 before the break	0
ETH CO2, ETH Energy, ETH Energy4, ETH Hashing1, ETH Hashing2 before the break	0
ETH CO2, ETH Energy, ETH Energy4, ETH Hashing1, ETH Hashing2, ETH Price before the break	0

ETH CO2, ETH Energy after the break	0

Note. BTC stands for Bitcoin, and ETH for Ethereum. CO2 relates to carbon emission. The log difference represents the change in the CO2 levels. The number represents the polynomials. Before and after break refers to the structural break.

Table 27A BIC minimisation for ETH

Lag	AIC	HQIC	SBIC
0	4.697	4.703	4.713
1	-3.811	-3.780	-3.730
2	-4.265	-4.210	-4.120*
3	-4.307	-4.228	-4.098
4	-4.343	-4.240	-4.069
5	-4.395	-4.268	-4.057
6	-4.420*	-4.269*	-4.018

Note. *optimal lag, AIC stands for Akaike Information Criterion, HQIC for Hannan-Quinn Schwarz's Information Criterion, and SBIC for Schwarz's Bayesian Information Criterion.

Table 28A IV correlation matrix for Ethereum

Variables	Residual
Residuals	1
Hashing power ETH	0.014
E-waste	0.002
Price ETH	-0.013
Energy BTC	-0.012
Hashing power BTC	0.023
Price BTC	-0.001

Note. ETH is Ethereum, energy refers to energy consumption. The variables represent log differences (change).

Table 29A The results for the unit root tests of the residuals ETH

Test name	P-value	Test statistic
Dickey–Fuller	0.000	-
Phillips–Perron	0.000	-
KPSS	-	0.152 (lag 18)

Note. KPSS stands for Kwiatkowski-Phillips-Schmidt-Shin. The critical value for the KPSS test with a 5% confidence level is 0.146.

Table 30A Cointegration test ETH

Rank	Rank<=(r+1)	Rank<=(p=3)
0	1318.834	2881.156
1	885.579	1562.322
2	399.120	676.743
3	277.623	277.623

Note. The null-hypothesis of non-cointegration can be rejected if the trace statistics value is higher than the max-lambda in rank =>1. The max-lambda is Rank<=(r+1). The trace statistics values are presented in the third column.

Table 31A BIC minimisation for BTC

Lag	AIC	HQIC	SBIC
0	-13.552	-13.547	-13.540
1	-13.797	-13.779	-13.748
2	-13.917	-13.885	-13.832
3	-13.964	-13.919	-13.843
4	-14.052	-13.993	-13.895
5	-14.2350	-14.162	-14.042
6	-14.282	-14.196	-14.052*
7	-14.306*	-14.206*	-14.040
8	-14.305	-14.191	-14.002
9	-14.296	-14.168	-13.957
10	-14.285	-14.144	-13.910

Note. *optimal lag, AIC stands for Akaike Information Criterion, HQIC for Hannan-Quinn Schwarz's Information Criterion, and SBIC for Schwarz's Bayesian Information Criterion.

Table 32A IV correlation matrix for Bitcoin

Variables	Residual
Residuals	1
Energy ETH	0.011
Hashing power ETH	-0.005
E-waste	0.051
Price ETH	0.015
Hasing power BTC	0.003
Price BTC	0.036

Note. BTC is Bitcoin, energy refers to energy consumption. The variables represent log differences (change).

Table 33A Results unit root test of the residuals BTC

Test name	P-value	Test statistic
Dickey–Fuller	0.000	-
Phillips–Perron	0.000	-
KPSS	-	0.203 (lag 0)

Note. KPSS stands for Kwiatkowski-Phillips-Schmidt-Shin. The critical value for the KPSS test is 0.146.

Table 34A Cointegration test BTC

Rank	Rank<=(r+1)	Rank<=(p=3)
0	769.541	1129.686
1	197.003	360.145
2	163.142	163.142

Note. The null-hypothesis of non-cointegration can be rejected if the trace statistics value is higher than the max-lambda in rank =>1. The max-lambda is Rank<=(r+1). The trace statistics values are presented in the third column.

Table 35 A – BIC minimisation for the ecosystem

Lag	AIC	HQIC	SBIC
0	-2.931	-2.926	-2.919

1	-3.220	-3.201	-3.171
2	-3.368	-3.337	-3.284
3	-3.436	-3.390	-3.315
4	-3.542	-3.483	-3.385
5	-3.730	-3.658	-3.537
6	-3.769	-3.682*	-3.539*
7	-3.780	-3.680	-3.514
8	-3.780	-3.666	-3.477
9	-3.782*	-3.655	-3.443
10	-3.774	-3.634	-3.400

Note. The optimal lag is indicated with an *, AIC stands for Akaike Information Criterion, HQIC for Hannan-Quinn Schwarz's Information Criterion, and SBIC for Schwarz's Bayesian Information Criterion.

Table 36A IV correlation matrix for the ecosystem

Variables	Residual
Residuals	1
Hashing power ETH	0.055
E-waste	0.005
Price ETH	-0.041
BTC hashing power	0.022
Price BTC	0.005

Note. ETH is Ethereum, BTC stands for Bitcoin, energy refers to energy consumption. The variables represent log differences (change).

Table 37A Results unit root test of the residuals for the ecosystem

Test name	P-value	Test statistic
Dickey–Fuller	0.000	-
Phillips–Perron	0.000	-
KPSS	-	0.119 (lag 22)

Note. KPSS stands for Kwiatkowski-Phillips-Schmidt-Shin. The critical value for the KPSS test is 0.146.

Table 38A Cointegration test for the ecosystem

Rank	Rank<=(r+1)	Rank<=(p=3)
0	728.936	1267.712
1	359.918	538.776
2	178.858	178.858

Note. The null-hypothesis of non-cointegration can be rejected if the trace statistics value is higher than the max-lambda in rank =>1. The max-lambda is Rank<=(r+1). The trace statistics values are presented in the third column.

Table 39A BIC minimisation for ETH before the break

Lag	AIC	HQIC	SBIC
0	-34.993	-34.983	-34.967
1	-35.175	-35.108	-34.996*
2	-35.254	-35.129*	-35.129*
3	-35.294	-35.111*	-34.808*
4	-35.345	-35.104	-34.706

5	-35.407*	-35.108	-34.613
6	-35.404	-35.047	-34.457

Note. The optimal lag is indicated with an *, AIC stands for Akaike Information Criterion, HQIC for Hannan-Quinn Schwarz's Information Criterion, and SBIC for Schwarz's Bayesian Information Criterion.

Table 40A IV correlation matrix for Ethereum before the break

Variables	Residual
Residuals	1
E-waste	0.051
Hashing power BTC	0.010
Price BTC	-0.069

Note. ETH is Ethereum, BTC is Bitcoin, energy refers to energy consumption. The variables represent log differences (change).

Table 41A Results unit root test of the residuals for ETH before the break

Test name	P-value	Test statistic
Dickey–Fuller	0.000	-
Phillips–Perron	0.000	-
KPSS	-	0.088 (lag 22)

Note. Note. KPSS stands for Kwiatkowski-Phillips-Schmidt-Shin. The critical value for the KPSS test is 0.146.

Table 42A Cointegration test for Ethereum before the break

Rank	Rank<=(r+1)	Rank<=(p=2)
0	103.050	20.409
1	123.459	20.409

Note. The null-hypothesis of non-cointegration can be rejected if the trace statistics value is higher than the max-lambda in rank =>1. The max-lambda is Rank<=(r+1). The trace statistics values are presented in the third column.

Table 43A BIC minimisation for ETH after the break

Lag	AIC	HQIC	SBIC
0	6.716	6.738	6.771
1	4.514	4.581	4.680
2	3.821	3.933	4.098
3	3.681*	3.838*	4.070*
4	3.702	3.904	4.202
5	3.774	4.021	4.385
6	3.716	4.007	4.438

Note. The optimal lag is indicated with an *, AIC stands for Akaike Information Criterion, HQIC for Hannan-Quinn Schwarz's Information Criterion, and SBIC for Schwarz's Bayesian Information Criterion.

Table 44A IV correlation matrix for Ethereum after the break

Variables	Residual
Residual	1
Hasing power ETH	0.128
e-waste	0.005
Price eth	0.026
Hasing power BTC	0.017

Price BTC	-0.001

Note. ETH is Ethereum, energy refers to energy consumption. The variables represent log differences (change).

Table 45A - Results unit root test of the residuals of ETH after the break

Test name	P-value	Test statistic
Dickey–Fuller	0.000	-
Phillips–Perron	0.000	-
KPSS	-	0.070 (lag 22)

Note. KPSS stands for Kwiatkowski-Phillips-Schmidt-Shin. The critical value for the KPSS test is 0.146.

Table 46A – Cointegration test for Eth after the break

Rank	Rank<=(r+1)	Rank<=(p=2)
0	103.050	123.459
1	20.409	20.419

Note. The null-hypothesis of non-cointegration can be rejected if the trace statistics value is higher than the max-lambda in rank =>1. The max-lambda is Rank<=(r+1). The trace statistics values are presented in the third column.

Table 47A - BIC minimisation for ecosystem considering the structural break

Lag	AIC	HQIC	SBIC
0	-4.912	-4.907	-4.899
1	-5.070	-5.052	-5.022
2	-5.142	-5.111*	-5.058*
3	-5.148	-5.103	-5.027
4	-5.149	-5.089	-4.992
5	-5.176*	-5.103	-4.983
6	-5.171	-5.085	-4.941

Note. The optimal lag is indicated with an *, AIC stands for Akaike Information Criterion, HQIC for Hannan-Quinn Schwarz's Information Criterion, and SBIC for Schwarz's Bayesian Information Criterion.

Table 48A – Results unit root test for the ecosystem considering the structural break

Test name	P-value	Test statistic
Dickey–Fuller	0.000	-
Phillips–Perron	0.000	-
KPSS	-	0.108 (lag 11)

Note. KPSS stands for Kwiatkowski-Phillips-Schmidt-Shin. The critical value for the KPSS test is 0.146.

Table 49A Cointegration test for ecosystem considering the structural break

Rank	Rank<=(r+1)	Rank<=(p=3)
0	879.157	1669.451
1	790.026	790.294
2	0.266	0.266

Note. The null-hypothesis of non-cointegration can be rejected if the trace statistics value is higher than the max-lambda in rank =>1. The max-lambda is Rank<=(r+1). The trace statistics values are presented in the third column.

Table 50A Results two-sample T-test for forecasted values

Null hypothesis	P-value
CO2 Ecosystem forecasted < Ecosystem CO2	0.691
CO2 Ecosystem forecasted = CO2 Ecosystem	0.617
CO2 Ecosystem forecasted > CO2 Ecosystem	0.308
CO2 ETH forecasted < CO2 ETH	0.525
CO2 ETH forecasted = CO2 ETH	0.948
CO2 ETH forecasted > CO2 ETH	0.473
CO2 BTC forecasted < CO2 BTC	0.487
CO2 BTC forecasted = CO2 BTC	0.975
CO2 BTC forecasted > CO2 BTC	0.512

Note. CO2 relates to carbon emission. BTC stands for Bitcoin, ETH refers to Ethereum.

APPENDIX B FIGURES

In this appendix, I present fourteen figures that did not appear in-text. For all clarity, I first list the figures in this part of the appendix.

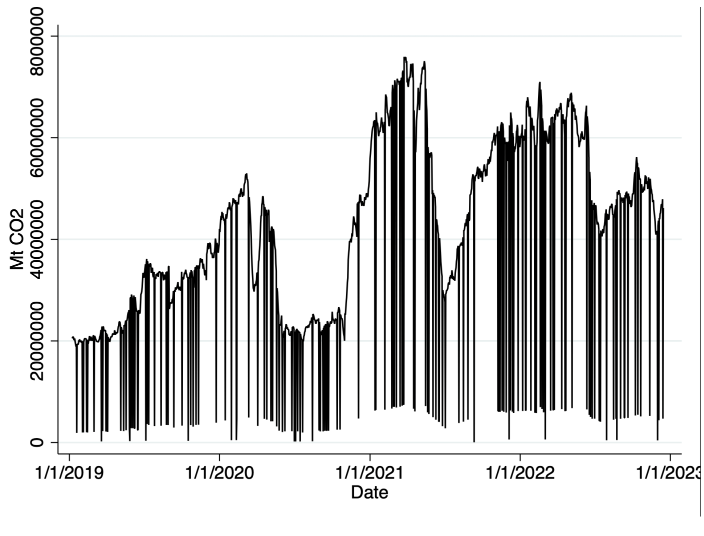

Figure 1B Plots before carbon BTC

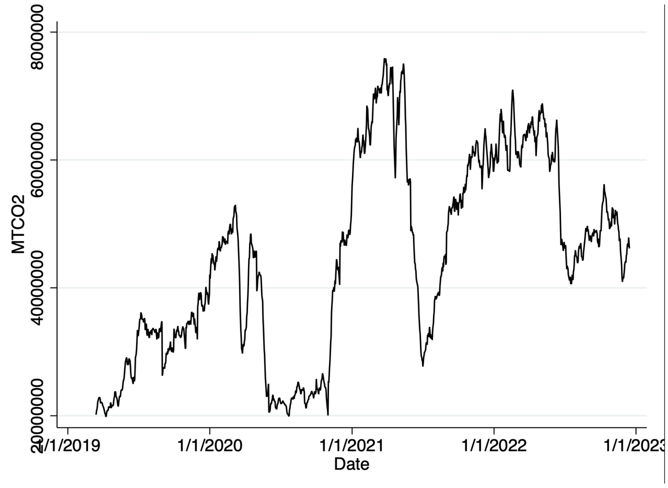

Figure 2B plots after treatment CO2 BTC

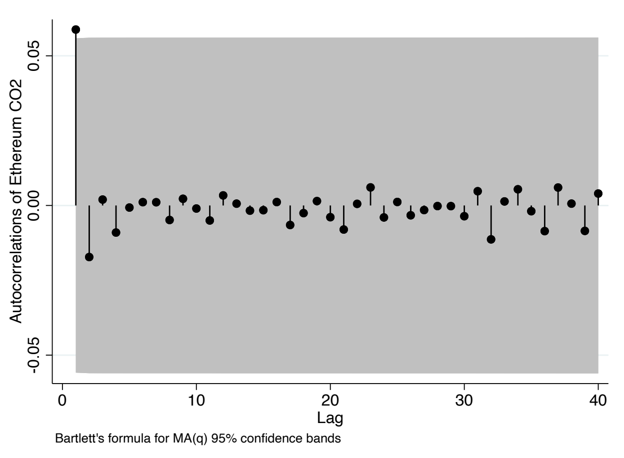

Figure 3B Autocorrelations of Ethereum's CO2

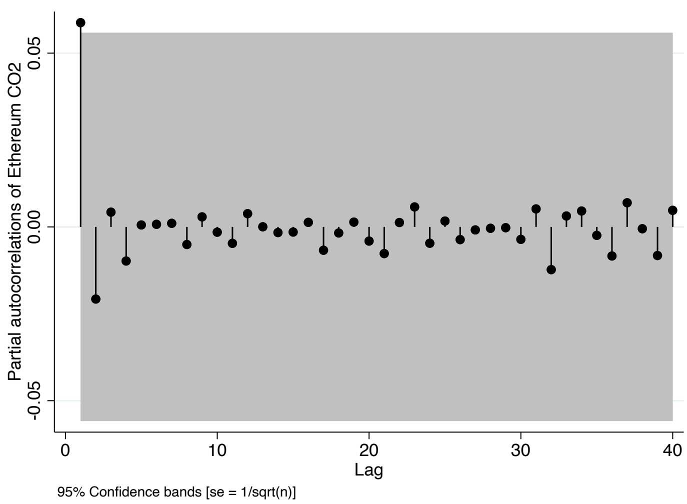

Figure 4B Partial autocorrelations of Ethereum's CO2

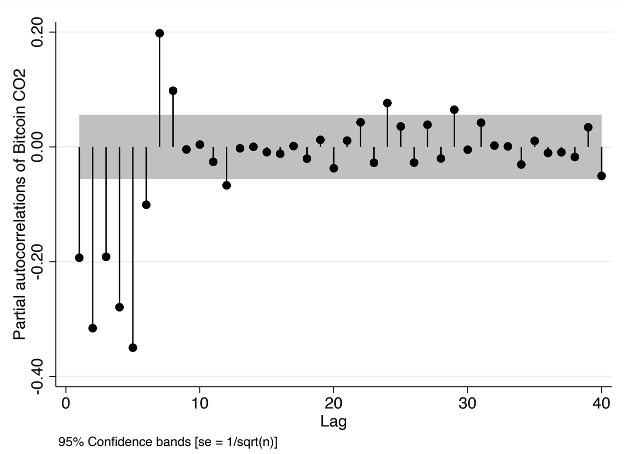

Figure 5B Autocorrelations of Bitcoin's CO2

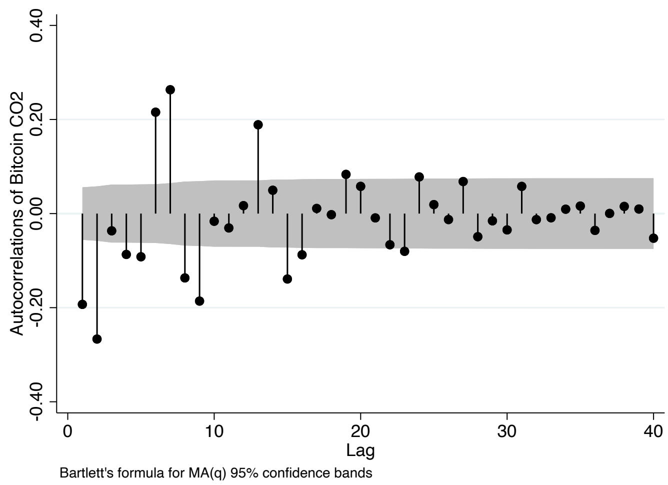

Figure 6B Partial autocorrelations of Bitcoin's CO2

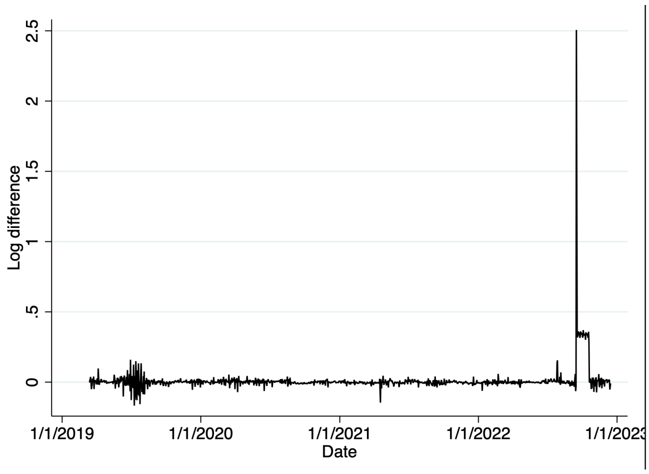

Figure 7B Plot of the ecosystem's benchmark

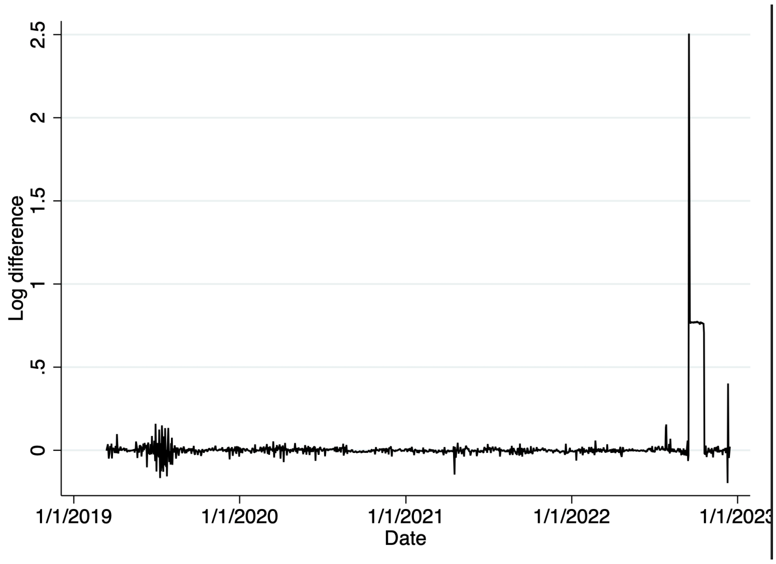

Figure 8B Plot of Ethereum's benchmark

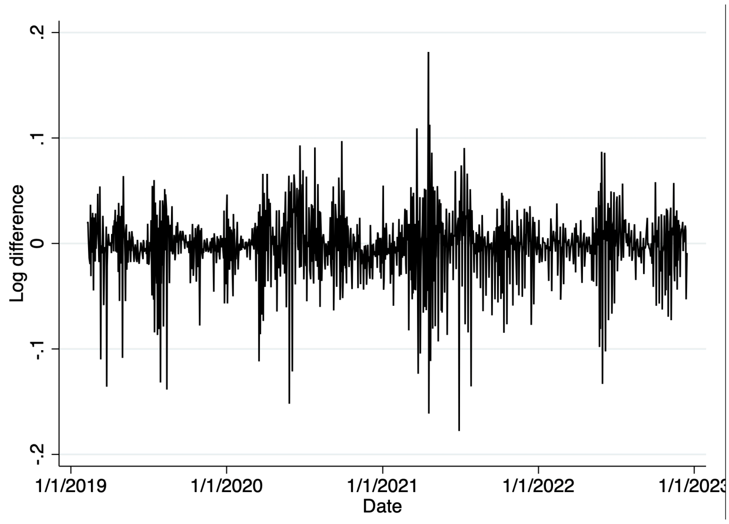

Figure 9B Plot of Bitcoin's benchmark

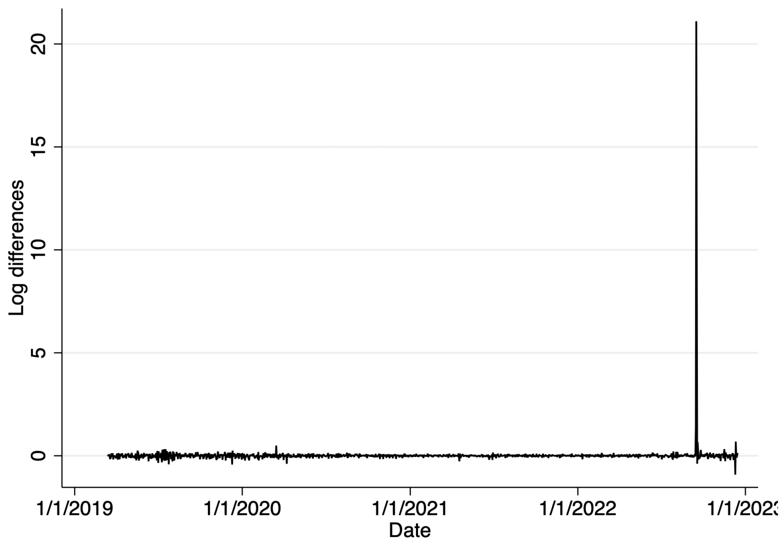

Figure 10B Plot with estimated carbon footprint per transaction Ethereum

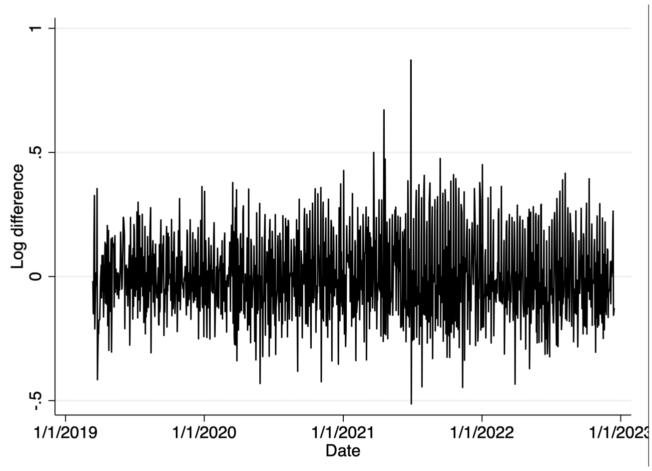

Figure 11B Plot with estimated carbon footprint per transaction Bitcoin

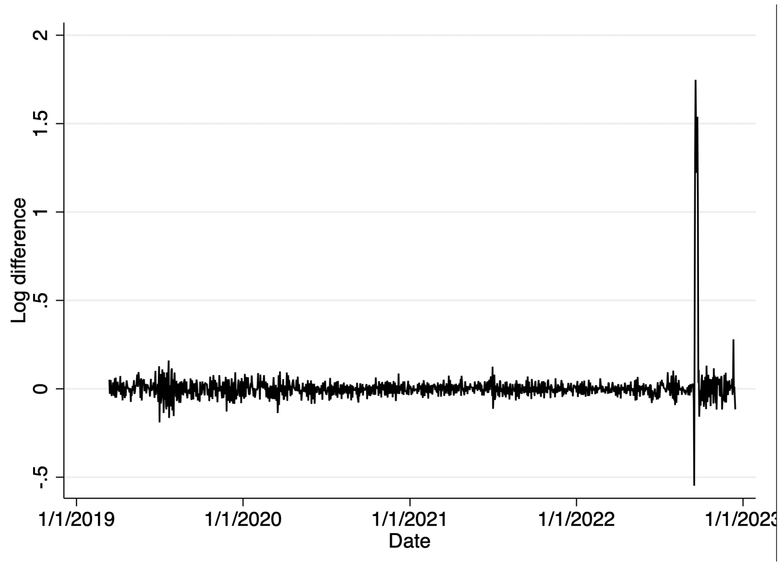

Figure 12B Plot with estimated carbon footprint per transaction the ecosystem

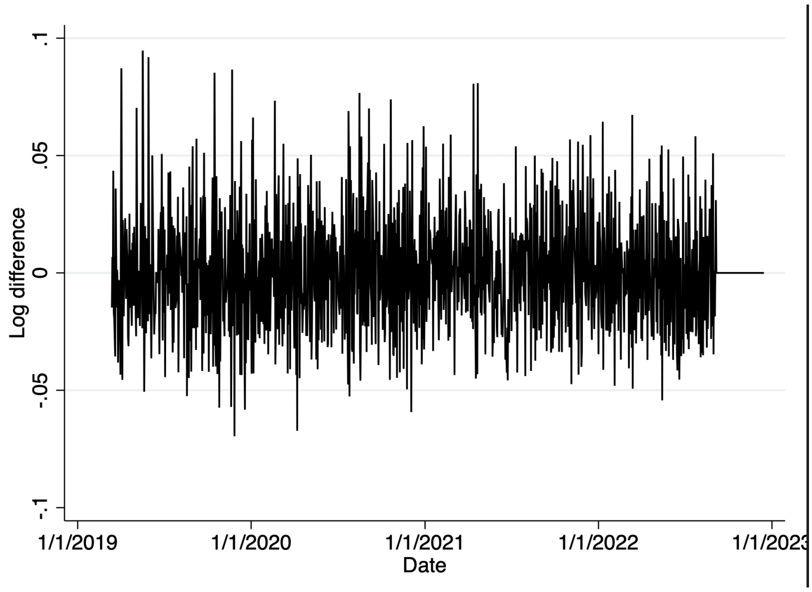

Figure 13B Plot of Ethereum considering the structural break

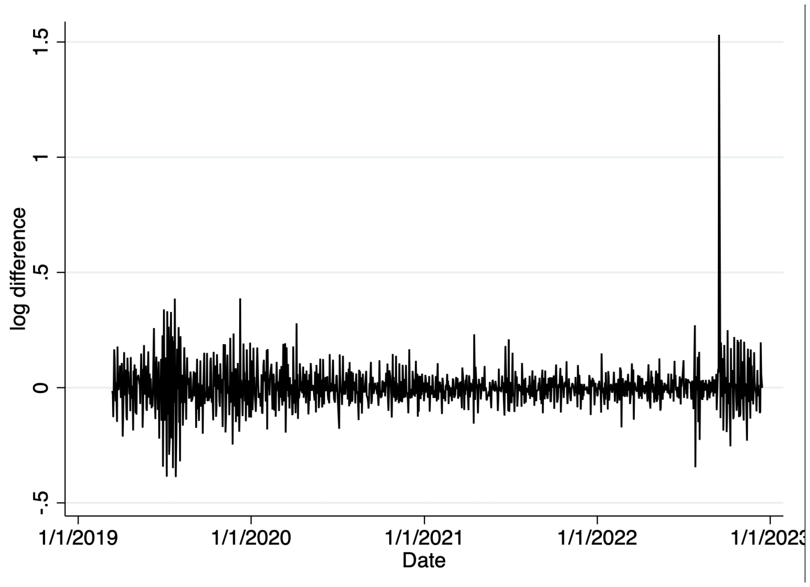

Figure 14B Plot of the ecosystem considering the structural break

APPENDIX C ABBREVIATIONS

Auto-Regressive (AR)

Auto Regressive Distributed Lag (ARDL)

Dickey-Fuller (DF)

Distributed Ledger Technology (DLT)

Diebold-Mariano (DM)

Instrumental Variable (IV)

Impulse Response Function (IRF)

Kwiatkowski-Phillips-Schmidt-Shin test (KPSS)

Newey West (NW)

Oxford Blockchain Strategy Framework (OBSF)

Proof of Stake (PoS)

Proof of Work (PoW)

Phillips-Perron (PP)

Standard Deviation (SD)

Vector Auto-Regressive (VAR)

Vector Error Correction Model (VECM)

Quandt-Likelihood Ratio (QLR)

Two-Stage-Least Square method (2SLS)

APPENDIX D CODE

```
////Installing packages
//ssc install ardl
//ssc install kpss
//ssc install johans

clear

//importing data

import excel "/Users/acmar1/Library/CloudStorage/OneDrive-
Personal/Desktop/Master/thesis/Data/master2.0.xlsx", sheet("Sheet2") firstrow

//////////////////////////////////////////////////////////////////////////
//fixing DATA
//////////////////////////////////////////////////////////////////////////

//drop missing values
drop if EstimatedMtCO2e == .

//twoway line EstimatedMtCO2e DATE

//distring
destring EwastektperYear, generate(btcewast) float
destring EstimatedTWhperYear, generate(energyeeth) float
destring etchashrate, generate(ethhash) float

destring bta_hash_rate, generate(btchash)
destring btc_price, generate(btcprice)

//drop string

drop EwastektperYear
drop EstimatedTWhperYear
drop etchashrate

drop bta_hash_rate
drop btc_price

//c02 per transaction
gen btcco2t =EstimatedMtCO2e /btc_transactions
gen ethco2t=Carbonemmisionktco2/transaction

//energy per transaction

gen btcenergyt = btcenergy /btc_transactions
gen ethenergyt =  energyeeth / transaction
```

```
//ecosystem carbon emission
gen co2ecosystem = btcco2t + ethco2t


////////////////////////////////////////////////////////////////////////
// summary STATS Raw data
////////////////////////////////////////////////////////////////////////

estpost sum btcco2t btchash btcprice btcewast protocol btcenergyt ethhash ethco2t price
ethenergyt protocol1 co2ecosystem, de
esttab using "stats__raw_data.csv" , cells ("mean(fmt(3)) p50(fmt(3))  sd(fmt(3)) max(fmt(3))
min(fmt(3)) skewness(fmt(3)) kurtosis(fmt(3)) count(fmt(3)) ") replace



//take out outliners//


twoway line EstimatedMtCO2e DATE

drop if EstimatedMtCO2e < 10000000

twoway line EstimatedMtCO2e DATE



// time variable
gen t = _n
tsset t

//log carbon emission
gen co2ecolog = log(co2ecosystem)
gen dco2ecolog = L1.co2ecolog - co2ecolog



//log difference ETH
gen co2log = log(ethco2t)
gen hashlog = log(ethhash)
gen pricelog = log(price)
gen ewastlog = log(btcewast)
gen energylog = log(ethenergyt)


gen dco2log = L1.co2log - co2log
```

```
gen dhashlog = L1.hashlog- hashlog
gen dpricelog = L1.pricelog- pricelog
gen dewastlog = L1.ewastlog- ewastlog
gen denergylog = L1.energylog-  energylog



replace dhashlog = 0 if missing(dhashlog)

//High polynomial variables ETH
gen denergylog2 =  denergylog^2
gen denergylog3 =  denergylog^3
gen denergylog4 =  denergylog^4

gen dhashlog2 =  dhashlog^2
gen dhashlog3 =  dhashlog^3
gen dhashlog4 =  dhashlog^4

gen dpricelog2 =  dpricelog^2
gen dpricelog3 =  dpricelog^3
gen dpricelog4 =  dpricelog^4

gen dewastlog2 =  dewastlog^2
gen dewastlog3 =  dewastlog^3
gen dewastlog4 =  dewastlog^4



//log difference BTC
gen btco2log = log(btcco2t)
gen btchashlog = log(btchash)
gen btcpricelog = log(btcprice)
gen btcewastlog = log(btcewast)
gen btcenergylog = log(btcenergyt)

gen dbtco2log = L1.btco2log - btco2log
gen dbtchashlog = L1.btchashlog - btchashlog
gen dbtcpricelog = L1.btcpricelog - btcpricelog
gen dbtcewastlog = L1.btcewastlog - btcewastlog
gen dbtcenergylog = L1.btcenergylog -  btcenergylog



//High order variables BTC
gen dbtcenergylog2 =  dbtcenergylog^2
gen dbtcenergylog3 =  dbtcenergylog^3
gen dbtcenergylog4 =  dbtcenergylog^4

gen dbtchashlog2 =  dbtchashlog^2
```

```
gen dbtchashlog3 =  dbtchashlog^3
gen dbtchashlog4 =  dbtchashlog^4

gen dbtcpricelog2 =  dbtcpricelog^2
gen dbtcpricelog3 =  dbtcpricelog^3
gen dbtcpricelog4 =  dbtcpricelog^4

gen dbtcewastlog2 =  dbtcewastlog^2
gen dbtcewastlog3 =  dbtcewastlog^3
gen dbtcewastlog4 =  dbtcewastlog^4




//////////////////////////////////////////////////////////////////////////////
// summary STATS without outliners
//////////////////////////////////////////////////////////////////////////////



//summary data

estpost sum btcco2t btchash btcprice btcewast  btcenergyt ethhash ethco2t price ethenergyt
protocol1 co2ecosystem, de
esttab using "stats_cleaned_data.csv" , cells ("mean(fmt(3)) p50(fmt(3))  sd(fmt(3))
max(fmt(3)) min(fmt(3)) skewness(fmt(3)) kurtosis(fmt(3)) count(fmt(3)) ") replace

///table above

//////////////////////////////////////////////////////////////////////////////
//  STATS test
//////////////////////////////////////////////////////////////////////////////

//ETH

//dickey-fuller
dfuller ethhash
dfuller ethco2t
dfuller price
dfuller ethenergyt
dfuller btcewast



//dickey-fuller log diff

dfuller dco2log
dfuller dhashlog
dfuller dpricelog
```

```
dfuller dewastlog
dfuller denergylog

//Phillips-Perron (PP) Test:

pperron ethhash
pperron ethco2t
pperron price
pperron ethenergyt
pperron btcewast



pperron co2log
pperron hashlog
pperron pricelog
pperron ewastlog
pperron energylog



pperron dco2log
pperron dhashlog
pperron dpricelog
pperron dewastlog
pperron denergylog

//Kwiatkowski-Phillips-Schmidt-Shin (KPSS) Test:
kpss ethhash
kpss ethco2t
kpss price
kpss ethenergyt
kpss btcewast


kpss dco2log
kpss dhashlog
kpss dpricelog
kpss dewastlog
kpss denergylog


//pac

pac dco2log
//pac dhashlog
//pac dpricelog
//pac dewastlog
//pac denergylog
```

```
//ac

ac dco2log
//ac dhashlog
//ac dpricelog
//ac dewastlog
//ac denergylog

////////////
//btc
////////////


//dickey-fuller
dfuller btcco2t
dfuller btchash
dfuller btcprice
dfuller btcenergyt



//Phillips-Perron (PP) Test:

pperron btcco2t
pperron btchash
pperron btcprice
pperron btcenergyt

//Kwiatkowski-Phillips-Schmidt-Shin (KPSS) Test:
kpss  btcco2t
kpss btchash
kpss btcprice
kpss btcenergyt



//dickey-fuller log

dfuller dbtco2log
dfuller dbtchashlog
dfuller dbtcpricelog
dfuller dbtcewastlog
dfuller dbtcenergylog

pperron dbtco2log
pperron dbtchashlog
pperron dbtcpricelog
pperron dbtcewastlog
pperron dbtcenergylog
```

```
kpss dbtco2log
kpss dbtchashlog
kpss dbtcpricelog
kpss dbtcewastlog
kpss dbtcenergylog

//pac

pac dbtco2log
//pac dbtchashlog
//pac dbtcpricelog
//pac dbtcewastlog
//pac dbtcenergylog


//ac

ac dbtco2log
//ac dbtchashlog
//ac dbtcpricelog
//ac dbtcewastlog
//ac dbtcenergylog

//ultimate variable
dfuller dco2ecolog
pperron dco2ecolog
kpss dco2ecolog

pac dco2ecolog
ac dco2ecolog


dfuller co2ecosystem
pperron co2ecosystem
kpss co2ecosystem

pac co2ecosystem
ac co2ecosystem


//summary log variables
//btc

estpost sum dbtco2log dbtchashlog dbtcpricelog dbtcewastlog dbtcenergylog dhashlog
dco2log dpricelog denergylog dco2ecolog,de
esttab using "stats_all_logs.csv" , cells ("mean(fmt(3)) p50(fmt(3))  sd(fmt(3)) max(fmt(3))
min(fmt(3)) skewness(fmt(3)) kurtosis(fmt(3)) count(fmt(3)) ") replace
```

```
///correlation matrix normal variables

cor btcco2t btchash btcprice btcewast btcenergyt

cor ethco2t ethhash price  btcewast ethenergyt


cor ethco2t ethhash price ethenergyt btcewast btcco2t btchash btcprice btcenergyt
putexcel set "correlation_matrix1.xlsx", sheet("Sheet1") modify
putexcel A1 = matrix(r(C))
putexcel close


cor dbtco2log dbtchashlog dbtcpricelog dbtcewastlog dbtcenergylog


////////////////////////////////////////////////////////////////////////////////
//  First Hyphotesis difference between BTC and ETH
////////////////////////////////////////////////////////////////////////////////

ttest dco2log == dbtco2log
ttest btcco2t == ethco2t

////////////////////////////////////////////////////////////////////////////////
//  second Hyphotesis
////////////////////////////////////////////////////////////////////////////////

//ETH

//energy (relevant)
eststo: reg dco2log denergylog
hettest
estat bgodfrey
eststo: newey dco2log denergylog, lag(1)

eststo: reg dco2log  denergylog denergylog2

hettest
estat bgodfrey
eststo: newey dco2log denergylog denergylog2 , lag(1)


eststo: reg dco2log denergylog denergylog2 denergylog3

hettest
estat bgodfrey
eststo: newey dco2log denergylog denergylog2 denergylog3 , lag(1)
```

```
eststo: reg dco2log denergylog denergylog2 denergylog3 denergylog4

esttab using Energy_eth.rtf, se r2 replace
eststo clear


//hash (relevant)
eststo: reg dco2log denergylog denergylog2 denergylog3   dhashlog
eststo: reg dco2log denergylog denergylog2 denergylog3   dhashlog2
eststo: reg dco2log denergylog denergylog2 denergylog3   dhashlog3
eststo: reg dco2log denergylog denergylog2 denergylog3   dhashlog4
esttab using hash_eth.rtf, se r2 replace
eststo clear

//ewast (not relevant)
eststo: reg dco2log denergylog denergylog2 denergylog3   dewastlog
eststo: reg dco2log denergylog denergylog2 denergylog3   dewastlog2
eststo: reg dco2log denergylog denergylog2 denergylog3   dewastlog3
eststo: reg dco2log denergylog denergylog2 denergylog3   dewastlog4
esttab using waste_eth.rtf, se r2 replace
eststo clear

//protocol (not relevant)
eststo: reg dco2log denergylog denergylog2 denergylog3   i.protocol1
esttab using protocol_eth.rtf, se r2 replace
eststo clear

//price (not relevant)
eststo: reg dco2log denergylog denergylog2 denergylog3   dpricelog
eststo: reg dco2log denergylog denergylog2 denergylog3   dpricelog2
eststo: reg dco2log denergylog denergylog2 denergylog3    dpricelog3
eststo: reg dco2log denergylog denergylog2 denergylog3    dpricelog4
esttab using price_eth.rtf, se r2 replace
eststo clear

////////////////////////////////////////////////////////////////////////////
//final formula
/////////////////////////////////////////////////////////////////////////////
eststo: reg dco2log denergylog denergylog2 denergylog3
hettest
estat bgodfrey
eststo: newey dco2log denergylog denergylog2 denergylog3, lag(1)
esttab using final_eth.rtf, se r2 replace
eststo clear



//residual stationarity
```

```
var dco2log denergylog denergylog2 denergylog3, lags(2)
predict resid, residuals


dfuller resid
pperron resid
kpss resid

//optimun amount of
varsoc dco2log denergylog denergylog2 denergylog3 , maxlag(6)
//var
var dco2log denergylog denergylog2 denergylog3 ,lag (1 2)

//Granger causality
vargranger


//cointegration test

johans dco2log denergylog denergylog2 denergylog3, lags(2)


//VECM
eststo clear
eststo: vec dco2log denergylog denergylog2 denergylog3, lags(2)
esttab using test_etf.rtf, se r2 replace
eststo clear


gen eth_carbon =.
vec dco2log denergylog denergylog2 denergylog3, lags(2)
predict temp
replace eth_carbon = temp
replace eth_carbon = 0 if missing(eth_carbon)


/////////////////////////////////////////////////////////////////
//impulse test /////////////////

varbasic dco2log denergylog denergylog2 denergylog3, lags(1 2) step(10) irf

//test for instrument variable

corr resid dhashlog dewastlog dpricelog dbtcenergylog dbtchashlog dbtcewastlog
dbtcpricelog

drop resid

//////////////////////////////////////////////////////////////////////////
//BTC
```

```
////////////////////////////////////////////////////////////////////////

//energy (relevant)
eststo: reg dbtco2log dbtcenergylog
hettest
estat bgodfrey
eststo: newey dbtco2log dbtcenergylog, lag(1)

eststo: reg dbtco2log dbtcenergylog dbtcenergylog2
hettest
estat bgodfrey
eststo: newey dbtco2log  dbtcenergylog dbtcenergylog2, lag(1)


eststo: reg dbtco2log dbtcenergylog  dbtcenergylog3
eststo: reg dbtco2log dbtcenergylog  dbtcenergylog4
hettest
estat bgodfrey
eststo: newey dbtco2log  dbtcenergylog dbtcenergylog4, lag(1)

esttab using Energy_btc.rtf, se r2 replace
eststo clear

//hash (not relevant)
eststo: reg dbtco2log dbtcenergylog  dbtcenergylog4 dbtchashlog
eststo: reg dbtco2log dbtcenergylog dbtcenergylog4 dbtchashlog2
eststo: reg dbtco2log dbtcenergylog  dbtcenergylog4 dbtchashlog3
eststo: reg dbtco2log dbtcenergylog dbtcenergylog4 dbtchashlog4
esttab using hash_btc.rtf, se r2 replace
eststo clear

//ewast (not relevant)
eststo: reg dbtco2log dbtcenergylog dbtcenergylog4 dbtcewastlog
hettest
estat bgodfrey
eststo: newey dbtco2log  dbtcenergylog dbtcenergylog4 dbtcewastlog, lag(1)

eststo: reg dbtco2log dbtcenergylog dbtcenergylog4 dbtcewastlog2
eststo: reg dbtco2log dbtcenergylog  dbtcenergylog4 dbtcewastlog3
eststo: reg dbtco2log dbtcenergylog dbtcenergylog4 dbtcewastlog4

esttab using waste_btc.rtf, se r2 replace
eststo clear


//price (not relevant)
eststo: reg dbtco2log dbtcenergylog dbtcenergylog4 dbtcpricelog
eststo: reg dbtco2log dbtcenergylog dbtcenergylog4 dbtcpricelog2
eststo: reg dbtco2log dbtcenergylog dbtcenergylog4 dbtcpricelog3
```

```
eststo: reg dbtco2log dbtcenergylog  dbtcenergylog4 dbtcpricelog4

esttab using price_btc.rtf, se r2 replace
eststo clear

///////////////////////////////////////////////////////////////////////
//final formula
///////////////////////////////////////////////////////////////////////
eststo: reg dbtco2log dbtcenergylog dbtcenergylog4
hettest
bgodfrey
eststo: newey dbtco2log dbtcenergylog dbtcenergylog4, lag(1)

esttab using final_btc.rtf, se r2 replace
eststo clear


//residual stationarity
var dbtco2log dbtcenergylog dbtcenergylog4, lag (1)

predict resid, residuals


dfuller resid
pperron resid
kpss resid

//optimun amount of
varsoc dbtco2log dbtcenergylog dbtcenergylog4, maxlag(6)
//var
var dbtco2log dbtcenergylog dbtcenergylog4, lag(1 2 3 4 5 6)

//Granger causality
vargranger



//cointegration test

johans dbtco2log dbtcenergylog dbtcenergylog4, lags(6)


//VECM

//VECM
eststo clear
eststo: vec dbtco2log dbtcenergylog dbtcenergylog4, lags(6) trend(t)
esttab using btc_test.rtf, se r2 replace
eststo clear
```

```
///////////////////////////////////////////////////////////////
//impulse test /////////////////

varbasic dbtco2log dbtcenergylog dbtcenergylog4, lags(1 2 3 4 5 6) step(10) irf


//test for instrument variable

corr resid denergylog dhashlog dewastlog dpricelog dbtchashlog  dbtcpricelog

drop resid


gen btc_carbon =.
vec dbtco2log dbtcenergylog dbtcenergylog4, lags(6) trend(t)
predict temp9
replace btc_carbon = temp9
replace btc_carbon = 0 if missing(btc_carbon)




///// gerental formula

gen co2_estimate_before =.
reg  dco2ecolog btc_carbon eth_carbon
hettest
bgodfrey
newey dco2ecolog btc_carbon eth_carbon, lag (1)

vec dco2ecolog btc_carbon eth_carbon, lags(6) trend(t)

predict temp80
replace co2_estimate_before = temp80




ttest co2_estimate_before == dco2ecolog


estpost sum co2_estimate_before btc_carbon eth_carbon,de
esttab using "stats_prediction_eco_second_logs.csv" , cells ("mean(fmt(8)) p50(fmt(8))
sd(fmt(8)) max(fmt(8)) min(fmt(8)) skewness(fmt(8)) kurtosis(fmt(8)) count(fmt(8)) ")
replace
```

```
//residual stationarity
var dco2ecolog btc_carbon eth_carbon, lag(1 2 3 4 5 6)

predict resid, residuals


dfuller resid
pperron resid
kpss resid

//optimun amount of
varsoc dco2ecolog btc_carbon eth_carbon, maxlag(10)
//var
var dco2ecolog btc_carbon eth_carbon, lag(1 2 3 4 5 6)

//Granger causality
vargranger

varbasic dco2ecolog btc_carbon eth_carbon, lags(1 2 3 4 5 6) step(10) irf


//cointegration test

johans dco2ecolog btc_carbon eth_carbon, lags(6)


//VECM
vec dco2ecolog btc_carbon eth_carbon, lags(6) trend(t)



////////////////////////////////////////////////////////////////
//impulse test /////////////////



corr resid dhashlog dewastlog dpricelog dbtchashlog  dbtcpricelog

drop resid

///////////////////////////////////////////////////////////////////////
//  third Hyphotesis
///////////////////////////////////////////////////////////////////////

//ETH


gen breaks= (t >1200)
gen breakx = breaks*denergylog
```

```
reg dco2log denergylog breakx breaks
reg dco2log denergylog denergylog2 denergylog3 breakx breaks
test breaks breakx
//Break found



reg dco2log denergylog
estat sbsingle



///////////////////////////////////////////////////////////////////////////
// 1200 (when it changed protocal)
///////////////////////////////////////////////////////////////////////////


//

//BTC

gen breakb= (t >724)
gen breakbx = breakb*dbtcenergylog

reg dbtco2log dbtcenergylog breakb breakbx
reg dbtco2log dbtcenergylog dbtcenergylog4 breakb breakbx
test breakb breakbx



reg dbtco2log dbtcenergylog
estat sbsingle

////// no structural break was found ////


//unit root rest

//dickey-fuller log diff

dfuller dco2log if t > 1200
dfuller dhashlog if t > 1200
dfuller dpricelog if t > 1200
dfuller dewastlog if t > 1200
dfuller denergylog if t > 1200

pperron dco2log if t > 1200
pperron dhashlog if t > 1200
pperron dpricelog if t > 1200
```

```
pperron dewastlog if t > 1200
pperron denergylog if t > 1200

kpss dco2log if t > 1200
kpss dhashlog if t > 1200
kpss dpricelog if t > 1200
kpss dewastlog if t > 1200
kpss denergylog if t > 1200

dfuller dco2log if t < 1200
dfuller dhashlog if t < 1200
dfuller dpricelog if t < 1200
dfuller dewastlog if t < 1200
dfuller denergylog if t < 1200


pperron dco2log if t < 1200
pperron dhashlog if t < 1200
pperron dpricelog if t < 1200
pperron dewastlog if t < 1200
pperron denergylog if t < 1200


kpss dco2log if t < 1200
kpss dhashlog if t < 1200
kpss dpricelog if t < 1200
kpss dewastlog if t < 1200
kpss denergylog if t < 1200

kpss dco2ecolog if t < 1200
kpss dco2ecolog if t > 1200
kpss co2ecosystem if t < 1200
kpss co2ecosystem if t > 1200

//dickey-fuller
dfuller ethhash if t > 1200
dfuller ethco2t if t > 1200
dfuller price if t > 1200
dfuller ethenergyt if t > 1200
dfuller btcewast if t > 1200

pperron ethhash if t > 1200
pperron ethco2t if t > 1200
pperron price if t > 1200
pperron ethenergyt if t > 1200
pperron btcewast if t > 1200

kpss ethhash if t > 1200
kpss ethco2t if t > 1200
kpss price if t > 1200
```

```
kpss ethenergyt if t > 1200
kpss btcewast if t > 1200

//dickey-fuller
dfuller ethhash if t < 1200
dfuller ethco2t if t < 1200
dfuller price if t < 1200
dfuller ethenergyt if t < 1200
dfuller btcewast if t < 1200

pperron ethhash if t < 1200
pperron ethco2t if t < 1200
pperron price if t < 1200
pperron ethenergyt if t < 1200
pperron btcewast if t < 1200

kpss ethhash if t < 1200
kpss ethco2t if t < 1200
kpss price if t < 1200
kpss ethenergyt if t < 1200
kpss btcewast if t < 1200




//////////////////////////////////////////////////////////////////////////
//  1st hyphotersis with break
//////////////////////////////////////////////////////////////////////////


ttest dco2log == dbtco2log if t < 1200
ttest dco2log == dbtco2log if t > 1200

ttest btcco2t == ethco2t if t < 1200
ttest btcco2t == ethco2t if t > 1200



//////////////////////////////////////////////////////////////////////////
//  2st hyphotersis with break
//////////////////////////////////////////////////////////////////////////


//ETH if t < 1200


//energy ( relevant)
eststo: reg dco2log denergylog if t < 1200
hettest
estat bgodfrey
```

```
eststo: newey dco2log denergylog if t < 1200  , lag(1)


eststo: reg dco2log denergylog denergylog2 if t < 1200
hettest
estat bgodfrey
eststo: newey dco2log denergylog denergylog2 if t < 1200  , lag(1)

eststo: reg dco2log denergylog denergylog3 if t < 1200
hettest
estat bgodfrey
eststo: newey dco2log denergylog  denergylog3 if t < 1200  , lag(1)


eststo: reg dco2log denergylog    denergylog4 if t < 1200
 hettest
estat bgodfrey
eststo: newey dco2log denergylog  denergylog4 if t < 1200  , lag(1)


esttab using Energy_eth.rtf, se r2 replace
eststo clear



//hash (relevant)
eststo: reg dco2log denergylog  denergylog4 dhashlog if t < 1200
hettest
estat bgodfrey

eststo: newey dco2log denergylog denergylog4 dhashlog if t < 1200  , lag(1)

eststo: reg dco2log denergylog denergylog4 dhashlog dhashlog2 if t < 1200
hettest
estat bgodfrey

eststo: newey dco2log denergylog denergylog4 dhashlog dhashlog2  if t < 1200  , lag(1)

eststo: reg dco2log denergylog denergylog4 dhashlog dhashlog2 dhashlog3 if t < 1200

eststo: reg dco2log denergylog denergylog4 dhashlog dhashlog2  dhashlog4 if t < 1200
esttab using hash_eth.rtf, se r2 replace
eststo clear

//ewast (not relevant)
eststo: reg dco2log denergylog denergylog4 dhashlog dhashlog2  dewastlog if t < 1200
eststo: reg dco2log denergylog denergylog4 dhashlog dhashlog2   dewastlog2 if t < 1200
eststo: reg dco2log denergylog denergylog4 dhashlog dhashlog2 dewastlog3 if t <1200
eststo: reg dco2log denergylog denergylog4 dhashlog dhashlog2 dewastlog4 if t <1200
esttab using waste_eth.rtf, se r2 replace
eststo clear
```

```
//protocol (not relevant)
eststo: reg dco2log denergylog denergylog4 dhashlog dhashlog2 i.protocol1 if t < 1200
esttab using protocol_eth.rtf, se r2 replace
eststo clear

//price ( relevant)
eststo: reg dco2log denergylog denergylog4 dhashlog dhashlog2   dpricelog if t < 1200
hettest
estat bgodfrey

eststo: newey dco2log denergylog denergylog4 dhashlog dhashlog2   dpricelog if t < 1200  ,
lag(1)

eststo: reg dco2log denergylog denergylog4 dhashlog dhashlog2   dpricelog dpricelog2 if t <
1200


eststo: reg dco2log denergylog denergylog4 dhashlog dhashlog2   dpricelog  dpricelog3 if t <
1200
eststo: reg dco2log denergylog denergylog4 dhashlog dhashlog2   dpricelog  dpricelog4 if t <
1200
esttab using price_eth.rtf, se r2 replace
eststo clear

//////////////////////////////////////////////////////////////////////////////
//final formula before break eth
//////////////////////////////////////////////////////////////////////////////
eststo: reg dco2log denergylog denergylog4 dhashlog dhashlog2   dpricelog   if t < 1200
eststo: newey dco2log denergylog denergylog4 dhashlog dhashlog2   dpricelog if t < 1200  ,
lag(1)

esttab using final_eth.rtf, se r2 replace
eststo clear



//var

varsoc dco2log denergylog denergylog4 dhashlog dhashlog2  dpricelog  if t < 1200 ,
maxlag(6)

var dco2log denergylog denergylog4 dhashlog dhashlog2  dpricelog  if t < 1200 ,lag(1)

//granger
vargranger

//cointegration
johans dco2log denergylog denergylog4 dhashlog dhashlog2 dpricelog if t < 1200, lags(1)
```

```
//test for instrument variable
var dco2log denergylog denergylog4 dhashlog dhashlog2  dpricelog     dpricelog if t < 1200,
lag(1)

predict resid, residuals

dfuller resid if t < 1200
pperron resid if t < 1200
kpss resid if t < 1200


corr resid  dewastlog dbtchashlog dbtcewastlog dbtcpricelog if t < 1200

drop resid

varbasic dco2log denergylog denergylog4 dhashlog dhashlog2  dpricelog  if t < 1200, lags(1)
step(10) irf

//
vec dco2log denergylog denergylog4 dhashlog dhashlog2  dpricelog if t < 1200, lags(1)


gen eth_carbon_before_break_fork =.
vec dco2log denergylog denergylog4 dhashlog dhashlog2 dpricelog if t < 1200, lags(1)
predict temp99 if t <1200
replace eth_carbon_before_break_fork= temp99


//after the break
//ETH if t > 1200


//energy (not relevant)
eststo: reg dco2log denergylog if t > 1200
eststo: newey dco2log denergylog if t >1200  , lag(1)
eststo: reg dco2log denergylog denergylog2 if t > 1200
eststo: reg dco2log  denergylog denergylog3 if t > 1200
eststo: reg dco2log denergylog  denergylog4 if t > 1200
esttab using Energy_eth.rtf, se r2 replace
eststo clear


//hash (relevant)
eststo: reg dco2log denergylog dhashlog if t > 1200
eststo: reg dco2log denergylog dhashlog2 if t > 1200
eststo: reg dco2log denergylog dhashlog3 if t >1200
eststo: reg dco2log denergylog dhashlog4 if t >1200
esttab using hash_eth.rtf, se r2 replace
eststo clear
```

```
//ewast (not relevant)
eststo: reg dco2log  denergylog  dewastlog if t > 1200
eststo: reg dco2log  denergylog  dewastlog2 if t > 1200
eststo: reg dco2log  denergylog dewastlog3 if t >1200
eststo: reg dco2log denergylog   dewastlog4 if t >1200
esttab using waste_eth.rtf, se r2 replace
eststo clear

//protocol (not relevant)
eststo: reg dco2log  denergylog  i.protocol1 if t > 1200
esttab using protocol_eth.rtf, se r2 replace
eststo clear

//price (not relevant)
eststo: reg dco2log   denergylog  dpricelog if t > 1200
eststo: reg dco2log  denergylog  dpricelog2 if t > 1200
eststo: reg dco2log  denergylog   dpricelog3 if t > 1200
eststo: reg dco2log  denergylog   dpricelog4 if t >1200
esttab using price_eth.rtf, se r2 replace
eststo clear

//////////////////////////////////////////////////////////////////////////
//Only Energy linear found sifnigicant nw
//////////////////////////////////////////////////////////////////////////

//best match
eststo: reg dco2log denergylog if t > 1200
hettest
estat bgodfrey

eststo: newey dco2log denergylog if t >1200  , lag(1)
esttab using final_eth.rtf, se r2 replace
eststo clear

varsoc dco2log denergylog  if t > 1200 , maxlag(6)

var dco2log denergylog if t > 1200 ,lag(1 2 3)
vargranger



//cointegration
johans dco2log denergylog if t > 1200, lags(3)

//test for instrument variable
newey dco2log denergylog if t > 1200, lag(1)

predict resid, residuals

dfuller resid if t > 1200
```

```
pperron resid if t > 1200
kpss resid if t > 1200


corr resid  dewastlog dbtchashlog dbtcewastlog dbtcpricelog if t > 1200

drop resid

varbasic dco2log denergylog  if t > 1200, lags(1 2 3) step(10) irf

//
vec dco2log denergylog if t > 1200, lags(3)


gen eth_carbon_after_break1 =.
vec dco2log denergylog if t > 1200, lags(3)
predict temp52 if t >1200
replace eth_carbon_after_break1 = temp52
replace eth_carbon_after_break1 = 0 if missing(eth_carbon_before_break)




//general formula


///// gerental formula

replace eth_carbon_before_break_fork = 0 if missing(eth_carbon_before_break_fork)
replace eth_carbon_after_break1 = 0 if missing(eth_carbon_after_break1)


gen ETH_co2_estimate1= eth_carbon_before_break_fork + eth_carbon_after_break1
replace ETH_co2_estimate1 = 0 if missing(ETH_co2_estimate1)



varsoc dco2ecolog btc_carbon ETH_co2_estimate1, maxlag(6)

eststo: var dco2ecolog btc_carbon ETH_co2_estimate1, lag (1 2)
eststo: vargranger


//cointegration
johans co2ecolog btc_carbon ETH_co2_estimate1 , lags(2)

//test for instrument variable
var dco2ecolog btc_carbon ETH_co2_estimate1, lag (1 2)

predict resid, residuals
```

```
dfuller resid if t > 1200
pperron resid if t > 1200
kpss resid if t > 1200


corr resid  dewastlog dbtchashlog dbtcewastlog dbtcpricelog

drop resid

varbasic co2ecolog btc_carbon ETH_co2_estimate1, lags(1 2) step(10) irf

//
vec dco2ecolog btc_carbon ETH_co2_estimate1, lags(2)



eststo clear



gen co2_estimate =.
eststo: reg dco2ecolog btc_carbon ETH_co2_estimate1
hettest
estat bgodfrey
eststo: newey dco2ecolog btc_carbon ETH_co2_estimate1, lag (1)
eststo: vec dco2ecolog btc_carbon ETH_co2_estimate1, lags(6)
esttab using mergedeco.rtf, se r2 replace
eststo clear

predict temp84
replace co2_estimate = temp84
///getting info on new variables

estpost sum dco2ecolog ETH_co2_estimate1 ,de
esttab using "stats_prediction_eco_breaksds_logs.csv" , cells ("mean(fmt(8)) p50(fmt(8))
sd(fmt(8)) max(fmt(8)) min(fmt(8)) skewness(fmt(8)) kurtosis(fmt(8)) count(fmt(8)) ")
replace



ttest dco2ecolog == co2_estimate




////////////////////////////////////////////////////////////////////////////
//  Predictive power
////////////////////////////////////////////////////////////////////////////
```

```
drop temp
drop temp52
drop temp80
drop temp84
drop temp9
drop temp99
//forecastes Btc


gen AR_Forecasts = .

forvalues s = 30/1290 {
	vec dbtco2log dbtcenergylog dbtcenergylog4 if t>=`s'-30 & t<`s', lags(6)
	predict temp if t==`s'
	replace AR_Forecasts = temp if t==`s'
	drop temp
}


gen MSE_MODEL_btc = sqrt((AR_Forecasts - dbtco2log)^2)



//////////////////////////////////////////////////////////////////////////
//banckmark btc
//////////////////////////////////////////////////////////////////////////



gen AR1_Forecasts = .

forvalues s = 30/1290 {
	reg dbtco2log L1.dbtco2log if t>=`s'-30 & t<`s'
	predict temp if t==`s'
	replace AR1_Forecasts = temp if t==`s'
	drop temp
}


gen MSE_BANCKMARK_btc = sqrt((AR1_Forecasts - dbtco2log)^2)




//forecastes eth before break

gen AR_eth_b_Forecasts = .
```

```
forvalues s = 30/1200 {
	vec dco2log denergylog denergylog4 dhashlog dhashlog2  dpricelog if t>=`s'-30 &
t<`s', lag(1)
	predict temp if t==`s'
	replace AR_eth_b_Forecasts = temp if t==`s'
	drop temp
}


gen MSE_MODEL_eth_b = sqrt((AR_eth_b_Forecasts - dco2log)^2)




///////////////////////////////////////////////////////////////////////////////
//banckmark before the break
///////////////////////////////////////////////////////////////////////////////


// AR --> Forecast before the break
gen AR_eth_bm_b_Forecasts = .

forvalues s = 30/1200 {
	newey dco2log L1.dco2log if t>=`s'-30 & t<`s', lag(1)
	predict temp if t==`s'
	replace AR_eth_bm_b_Forecasts = temp if t==`s'
	drop temp
}


gen MSE_BANCKMARK_eth_b = sqrt((AR_eth_bm_b_Forecasts - dco2log)^2)



/////////////////////////////////////////////////////////////
//ETH AFTER THE BREAK
/////////////////////////////////////////////////////////////
//reg dco2log  denergylog if t > 1200




//forecastes eth after break




gen AR_eth_a_Forecasts = .

forvalues s = 1200/1290 {
```

```
	vec  dco2log L1.denergylog  if t>=`s'-30 & t<`s', lag(3)
	predict temp if t==`s'
	replace AR_eth_a_Forecasts = temp if t==`s'
	drop temp
}


gen MSE_MODEL_eth_a = sqrt((AR_eth_a_Forecasts - dco2log)^2)



//////////////////////////////////////////////////////////////////////////
//banckmark after the break
//////////////////////////////////////////////////////////////////////////


// AR --> Forecast before the break
gen AR_eth_bm_a_Forecasts = .

forvalues s = 1200/1290 {
	newey dco2log L1.dco2log if t>=`s'-30 & t<`s', lag(1)
	predict temp if t==`s'
	replace AR_eth_bm_a_Forecasts = temp if t==`s'
	drop temp
}


gen MSE_BANCKMARK_eth_a = sqrt((AR_eth_bm_a_Forecasts - dco2log)^2)


///getting all together




replace AR_eth_a_Forecasts = 0 if missing(AR_eth_a_Forecasts)
replace AR_eth_b_Forecasts = 0 if missing(AR_eth_b_Forecasts)

replace MSE_BANCKMARK_eth_a = 0 if missing(MSE_BANCKMARK_eth_a)
replace MSE_BANCKMARK_eth_b = 0 if missing(MSE_BANCKMARK_eth_b)

replace MSE_MODEL_eth_a = 0 if missing(MSE_MODEL_eth_a)
replace MSE_MODEL_eth_b = 0 if missing(MSE_MODEL_eth_b)



gen co2_eth_total_final = AR_eth_a_Forecasts + AR_eth_b_Forecasts

gen MSE_MODEL_eth_final = MSE_MODEL_eth_a + MSE_MODEL_eth_b
```

```
gen MSE_BANCKMARK_eth_final = MSE_BANCKMARK_eth_a +
MSE_BANCKMARK_eth_b


gen co2_estimate_final_mf = .
forvalues s = 60/1290 {
	vec dco2ecolog co2_eth_total_final AR_Forecasts if t>=`s'-30 & t<`s', lag(3)
	predict temp if t==`s'
	replace co2_estimate_final_mf = temp if t==`s'
	drop temp
}



replace AR_eth_bm_a_Forecasts = 0 if missing(AR_eth_bm_a_Forecasts)
replace AR_eth_bm_b_Forecasts = 0 if missing(AR_eth_bm_b_Forecasts)

gen co2_eth_bm_total_final = AR_eth_bm_a_Forecasts + AR_eth_bm_b_Forecasts



gen co2_bm_final_mf = .
forvalues s = 60/1290 {
	reg dco2ecolog L1.dco2ecolog if t>=`s'-30 & t<`s'
	predict temp if t==`s'
	replace co2_bm_final_mf = temp if t==`s'
	drop temp
}



drop if co2_estimate_final_mf == .



estpost sum co2_bm_final_mf co2_eth_bm_total_final  AR1_Forecasts ,de
esttab using "stats_bm_final_logs.csv" , cells ("mean(fmt(3)) p50(fmt(3))  sd(fmt(3))
max(fmt(3)) min(fmt(3)) skewness(fmt(3)) kurtosis(fmt(3)) count(fmt(3)) ") replace


estpost sum co2_estimate_final_mf co2_eth_total_final   AR_Forecasts ,de
esttab using "stats_predi_final_mf_logs.csv" , cells ("mean(fmt(3)) p50(fmt(3))  sd(fmt(3))
max(fmt(3)) min(fmt(3)) skewness(fmt(3)) kurtosis(fmt(3)) count(fmt(3)) ") replace

//droping day of change in eth
```

```
//drop if co2_eth_bm_total_final >200
//


gen rmse_estimate = sqrt((co2_estimate_final_mf - dco2ecolog)^2)


gen rmse_estimate_bm = sqrt((co2_bm_final_mf - dco2ecolog)^2)

//diobold marino test RSME
ttest rmse_estimate == rmse_estimate_bm
ttest MSE_MODEL_eth_final == MSE_BANCKMARK_eth_final
ttest MSE_MODEL_btc == MSE_BANCKMARK_btc


//ttest estimate against actual


ttest co2_estimate_final_mf == dco2ecolog
ttest co2_eth_total_final ==  dco2log
ttest AR_Forecasts == dbtco2log




//plot benchmark


twoway line co2_eth_bm_total_final DATE //eth
twoway line AR1_Forecasts DATE // BTC
twoway line co2_bm_final_mf DATE //ECOSYSTEM



//plot estiamtion



twoway line eth_carbon DATE //eth
twoway line btc_carbon DATE // BTC
twoway line co2_estimate_before DATE //ECOSYSTEM



twoway line dco2ecolog DATE //ECOSYSTEM
twoway line ETH_co2_estimate1 DATE //eth
```

```
 //drop co2_estimate_before_break
 //drop co2_estimate_after_break
 //
// drop co2_bm_before
// drop co2_bm_after
//
// drop rmse_estimate_before
// drop rmse_estimate_after
//
// drop rmse_estimate_before_bm
// drop rmse_estimate_after_bm
//
//
//
//
```