sensors

Article

Decentralized IoT Data Authentication with Signature Aggregation

Jay Bojič Burgos * and Matevž Pustišek

Faculty of Electrical Engineering, University of Ljubljana, 1000 Ljubljana, Slovenia; matevz.pustisek@fe.uni-lj.si
* Correspondence: jay.bojicburgos@fe.uni-lj.si

Abstract: The rapid expansion of the Internet of Things (IoT) has introduced significant challenges in data authentication, necessitating a balance between scalability and security. Traditional approaches often rely on third parties, while blockchain-based solutions face computational and storage bottlenecks. Our novel framework employs edge aggregating servers and Ethereum Layer 2 rollups, offering a scalable and secure IoT data authentication solution that reduces the need for continuous, direct interaction between IoT devices and the blockchain. We utilize and compare the Nova and Risc0 proving systems for authenticating batches of IoT data by verifying signatures, ensuring data integrity and privacy. Notably, the Nova prover significantly outperforms Risc0 in proving and verification times; for instance, with 10 signatures, Nova takes 3.62 s compared to Risc0's 369 s, with this performance gap widening as the number of signatures in a batch increases. Our framework further enhances data verifiability and trust by recording essential information on L2 rollups, creating an immutable and transparent record of authentication. The use of Layer 2 rollups atop a permissionless blockchain like Ethereum effectively reduces on-chain storage costs by approximately 48 to 57 times compared to direct Ethereum use, addressing cost bottlenecks efficiently.

Keywords: blockchain; IoT; data authentication; signature aggregation; rollup; SNARK

Citation: Bojič Burgos, J.; Pustišek, M. Decentralized IoT Data Authentication with Signature Aggregation. *Sensors* **2024**, *24*, 1037. https://doi.org/10.3390/s24031037

Academic Editors: Wilfried Gappmair, Erich Leitgeb, Maja Matijašević and Mario Kusek

Received: 31 December 2023
Revised: 23 January 2024
Accepted: 1 February 2024
Published: 5 February 2024

1. Introduction

In the Internet of Things (IoT) realm, ensuring device identity and data authenticity is highly important. Data authenticity refers to the veracity and non-repudiation of data, ensuring that they have not been tampered with or forged. Data authentication and integrity are important for several reasons. Firstly, they safeguard the reliability and trustworthiness of data, ensuring that they accurately reflects the real-world conditions they represents. This is particularly critical in Industrial Internet of Things (IIoT) scenarios [1], where data are used to make informed decisions about critical assets, such as manufacturing equipment or power grids. Invalid or manipulated data can lead to erroneous actions, potentially causing disruptions, safety hazards, or financial losses.

The growing trend of data sharing and monetization within the IoT landscape underscores the need for robust data authenticity mechanisms. This is because companies and individuals increasingly exchange or purchase IoT data for various applications, such as predictive maintenance, smart home optimization, or healthcare analytics [2–4]. Data authenticity also plays a crucial role in supporting the development of artificial intelligence (AI) applications [5]. AI algorithms rely on vast amounts of high-quality data to train and refine their models. If the underlying data are not authentic, the AI models may generate erroneous or misleading insights, potentially leading to flawed decisions and actions.

Blockchain technology, a distributed ledger system known for its transparency, security, and immutability, offers a promising solution to address the challenges of IoT data authentication. By leveraging blockchains' inherent features, organizations can establish a trustless environment where IoT data can be securely shared and verified without intermediaries [6].

Blockchain-based IoT data authentication solutions can enable several valuable features, including the following:

- Data Integrity: Blockchains' tamper-proof nature ensures that IoT data remain unaltered throughout their lifecycle, preventing unauthorized modifications or data breaches.
- Device Authentication: A blockchain can serve as a secure repository for device identities, enabling reliable authentication of IoT devices and preventing unauthorized access or data injection from malicious or compromised devices.
- Provenance Tracking: Blockchains' ability to track data origin and movement provides traceability for IoT data, allowing for the verification of data sources and provenance.

However, while blockchains offer significant benefits for IoT data authentication, they also presents challenges that must be addressed. Scalability and cost are two significant issues that must be carefully considered [6]. The resource-constrained nature of many IoT devices raises concerns about the computational overhead and energy consumption associated with blockchain transactions. Moreover, the cost of maintaining and operating blockchain networks could be prohibitive for certain IoT applications.

The referenced paper [7] examines the application of permissioned blockchain architectures within a data authentication framework for IoT. Our stance is that relying solely on such architectures introduces limitations that could restrict the framework's scalability and broader applicability. While efficient and controllable, these systems risk segregating IoT networks from the dynamic and technologically rich broader blockchain ecosystem, which boasts considerable liquidity and a substantial user base. Additionally, private blockchains necessitate the establishment of validators, adding layers of complexity and an element of centralization, which runs counter to the ideal of a decentralized system. A fully blockchain-based data authentication solution for IoT devices based on public blockchains might also be impractical due to the prohibitive costs involved.

To address these challenges, our proposed framework takes a different approach from conventional blockchain-based authentication methods, which typically require direct interactions between resource-limited IoT devices and blockchain nodes [8]. We propose a layered architecture that decouples the generation and authentication of IoT data from the blockchain itself. In this model, IoT devices transmit their data to edge aggregating servers. These aggregators, acting as intermediary nodes and are equipped with ample computational resources to verify IoT signatures. Only the aggregated signatures/proofs or their hashes are relayed to the blockchain, serving as timestamps for the IoT signature-verification process.

This approach offloads the computational burden from IoT devices and creates a highly modular and scalable system. On the blockchain side, we employ scalable L2 blockchain scaling solutions, such as rollups, to address the issue of blockchain transaction throughput. By leveraging these scaling solutions, the blockchain can handle a significantly higher volume of IoT data without compromising security or performance. To further enhance scalability, edge aggregating servers, which collect and aggregate signatures from multiple IoT devices, are introduced. The increase in IoT devices and the scalability issues that these present can now be solved by introducing more aggregators and/or new rollups on the blockchain side.

- Our framework prioritizes data privacy by keeping sensitive IoT data confidential. Instead of allowing IoT devices to communicate directly with the blockchain, we handle all data verification through edge aggregator servers. This approach leverages Zero-Knowledge Proofs (ZKPs), enabling us to verify data without exposing the actual information, thus enhancing data privacy even further.

The implemented framework enabled us to conduct realistic experiments, providing detailed insights into the operation of a trusted, decentralized IoT solution. This solution stands out for its emphasis on data authenticity and the incorporation of Zero-Knowledge privacy principles. The key research questions investigated in this paper are as follows:

- Impact of different proving systems: Our exploration delves into various proving systems, examining their impact on the efficiency of proving and verification times and the sizes of the proofs required to ensure efficient data authenticity. Within our solution, we observed a shift in the performance bottleneck. It moves from the blockchain network to an off-chain system, which is particularly evident during the proving and verification stages.

- On-chain storage costs: The research also includes an assessment of the financial implications of using public blockchain networks for data storage. This component is vital to maintain a high level of decentralization and trust, providing a clear advantage over private or consortium blockchain networks. Our analysis balances the need for decentralization with the practical aspects of storage costs on public blockchains.

The contributions of our work can be summarized as follows:

- Development of a novel framework: We introduced a unique framework that integrates edge aggregating servers with Ethereum Layer 2 rollups. This design enhances scalability and security in IoT data authentication, minimizing the need for continuous direct interactions between IoT devices and the blockchain. By recording essential information on L2 rollups, our framework ensures data verifiability and trust, creating an immutable and transparent authentication record.
- Analysis of the viability of proving systems for our framework: Our work explored and compared the Nova and Risc0 proving systems, focusing on their efficiency in proof generation and verification times, as well as the proof sizes required for data authenticity. We discovered that employing proof recursion and compression is crucial for achieving superior performance in authentication, surpassing the efficiency of proving a single signature verification, and even outperforming methods like ECDSA batch verification.
- Cost-effective on-chain storage solutions: We assessed the financial implications of using public blockchain networks for data storage. The use of Layer 2 rollups led to significant reductions in on-chain storage costs. Our findings reveal that leveraging rollups atop public blockchains like Ethereum offers low fees while keeping us integrated within the larger, open-source public ecosystem, preventing isolation from the broader market.

The paper begins with an Introduction highlighting the importance of data authenticity in IoT and the potential of blockchain technology to address related challenges. Section 2, the Literature Review, overviews existing IoT data authentication approaches and briefly presents key fundamentals of our research as layered blockchain architectures and proving systems. Section 3 details the proposed solution, dividing it into off-chain and on-chain components. Section 4, Results, presents the findings from our tests on proving times, verification times, proof sizes, and on-chain storage costs. Section 5 discusses these results. Section 6 presents a set of conclusions summarizing the key insights and indicates possible future research.

2. Literature Review

Since our goal is to establish a framework for IoT data authentication that does not depend on permissioned blockchain networks or centralized solutions—which often entail reliance on third parties and carry inherent trust limitations—this section provides an overview of existing approaches to IoT data authentication. We also present the concept of layered blockchain structures and explore proving systems. These elements are crucial for the implementation of our proposed solution.

2.1. Existing Approaches to IoT Data Authentication

Data authentication plays an important role in IoT systems with some key challenges, including the following:

- Device authentication (registration and identity management) : Essential for preventing rogue device infiltration and ensuring data originate from verified sources, thereby maintaining IoT ecosystem integrity [7].

- Data integrity: IoT data must remain unchanged during transmission and storage to ensure their reliability [8].
- Data privacy: Protecting the confidentiality of sensitive information collected with IoT devices [9].
- Cyber security: IoT systems are vulnerable to attacks like DDoS, Sybil, and eavesdropping, which can disrupt operations and compromise data integrity [10,11].

Building on the understanding of these challenges, we find that while numerous solutions have been proposed, few manage to address all aspects of IoT data authentication in a decentralized manner.

The paper [8] proposes a solution for ensuring data integrity stored in cloud environments, particularly for Internet of Things (IoT) applications. The key idea is to use cryptographic hashes and smart contracts within a blockchain framework, utilizing Ethereum as the underlying blockchain. When data are stored in the cloud, a unique cryptographic hash of the data are generated and recorded on the blockchain. This hash serves as a digital fingerprint. Smart contracts are then used to verify the integrity of the data automatically. When data are retrieved, their hash is recalculated and compared with the hash stored on the blockchain. If the hashes match, it confirms that the data have not been altered, ensuring their integrity. While the approach presents a simple yet effective way to achieve data integrity, there are limitations to this approach. Not only does the use of Ethereum present considerable scalability issues since the costs would be too high, but the system focuses solely on verifying whether the data have been altered while in storage. It does not provide insights into the authenticity of the data themself. For instance, it does not verify whether the data originated from the correct IoT device or if they were authentic at the point of creation. This means that while the system can assure that the data have not been tampered with since they were stored, it cannot guarantee their initial authenticity or source. Therefore, additional mechanisms would be needed to validate the data's origin and authenticity before they are stored and hashed on the blockchain.

In contrast to this approach, another intriguing solution is presented in the paper [12], which discusses combining edge computing with a blockchain to enhance authentication in IoT networks. Edge servers are primarily responsible for authenticating IoT devices by verifying their credentials against blockchain-stored data, ensuring network access is limited to authorized devices only. Beyond authentication, these edge servers also contribute to efficient data management. Edge computing is crucial for real-time data processing in IoT applications, reducing latency and bandwidth demands. However, the system does place a certain level of trust in edge servers. While the blockchain ensures the integrity of the data it stores, the accuracy and security of the data initially provided by edge servers are vital. The system employs advanced cryptographic methods, including elliptic curve cryptography, to enhance security and protect against data compromise. Therefore, while blockchain technology ensures data immutability, the overall reliability and security of the system hinges on the trustworthiness and security measures implemented at the edge-server level.

A similar approach is the DIoTA framework [13], which offers a layered decentralized ledger architecture to enhance IoT data authenticity. It employs a unique edge–global structure, where each edge ledger serves specific IoT devices, and a global ledger interlinks these edge ledgers. This setup facilitates efficient cross-ledger data verification and incorporates a lightweight data authentication scheme to reduce the computational load on IoT devices. However, despite its innovative structure, DIoTA encounters significant limitations. Its reliance on a permissioned blockchain undermines the true essence of decentralization. The framework necessitates the establishment of validators, typically operated by entities within the system, such as IoT device owners or data analytic service providers. This model drifts away from a fully decentralized approach, as it centralizes control to a certain extent. If the number of these validators is limited or if they are heavily concentrated within a few entities, the system's security and decentralization are compromised.

On the positive side, DIoTA introduces some valuable concepts. For instance, the use of efficient data authentication methods for IoT communication with the ledger, specifically Hash-based message authentication code (HMAC) or Cypher-based message authentication code (CMAC), are noteworthy. This aspect could be integrated into our approach to enhance the efficiency of IoT device interactions, maintaining data integrity while minimizing resource utilization.

ZeroTrustBlock [14] presents a similar solution in the healthcare field that utilizes a permissioned blockchain based on a Hyperledger to enhance the security, privacy, and interoperability of sensitive medical data with existing approaches, akin to IoT data authentication.

Another paper [15] explores an approach to integrating blockchain technology with IoT data authentication, specifically focusing on smart grids and smart meters. This method incorporates Zero-Knowledge Proof (ZKP) to enhance anonymity and secure data within the blockchain environment. The solution involves using blockchain technology and Zero-Knowledge Proof to secure data from smart meters within a smart grid environment. This approach aims to prevent data counterfeiting and personal information infringement. The system utilizes Ethereum's smart contract functionality, incorporating Zero-Knowledge Proof to ensure data integrity and confidentiality. A significant drawback of this approach is that it requires IoT devices to write data directly onto the blockchain. This method is impractical and expensive due to the high financial costs of blockchain transactions.

While ZKP provides a method to verify the accuracy of information without revealing the data itself, the paper's approach seems more focused on post-storage data integrity rather than authenticating incoming data from IoT devices. There is a lack of a clear link between continuous data generation by IoT devices and the initial two values used during registration, potentially allowing fabricated data to be sent to the server without a way to verify its authenticity against the initially registered values.

In the context of enhancing security in IoT environments, Xu et al.'s study introduces a significant development with their Certificateless Aggregate Signature (CLAS) scheme [16], which is meant to be used for smart home applications. This scheme, addressing security concerns and the issue of private key leakage of similar schemes, operates in a certificateless environment. It aims to simplify key management and reduce computational overhead, especially in scenarios with numerous interconnected devices and substantial data flow. However, the inherent possession of individual keys in many IoT applications raises questions about the necessity and practicality of transitioning to a new scheme like CLAS.

In contrast, our approach, leveraging the widely used ECDSA in conjunction with zk-SNARKs, offers a more direct 'plug and play' solution, obviating the need to adopt a different scheme. Our method, using ZKPs, also stands out in terms of privacy preservation. It is nonetheless an interesting approach, since the use of a different schemes like CLAS actually allows the aggregation of signatures, while our approach creates a proof of signature verification as the aggregate.

The solution in [17] describes a security method for Internet of Things (IoT) networks that focuses on making sure each device (I-Node) can be trusted before it sends information through the network. It does this by setting up a system where devices have to prove their identity using a digital signature. The network selects certain devices (I-G Nodes) to act as checkpoints, collecting and verifying information from other devices. These checkpoints then send all the verified information to a central point (data center) for storage.

However, this system introduces unnecessary complexity by imposing verification duties on I-G Nodes without clear incentives or benefits for their participation. A more streamlined approach might involve edge servers that manage the computational tasks, thereby reducing the burden on individual IoT devices/nodes and simplifying the network's architecture. The proposed solution seems to occupy an awkward middle ground between decentralized and centralized systems, potentially inheriting the drawbacks of both without reaping the full benefits of either.

On the other hand, the work presented in [18] proposes a method to enhance privacy and computational efficiency in Industrial IoT systems. Utilizing the Paillier cryptosystem for privacy preservation and the Elliptic Curve Digital Signature Algorithm (ECDSA) with batch verification, the scheme addresses the critical need for secure data aggregation in edge-supported environments.

However, while this solution offers advancements in privacy protection, it also introduces additional computational work beyond ECDSA signing that may be prohibitive for IoT devices with limited resources. Furthermore, the use of batch ECDSA verification, as presented in [19], introduces a novel approach to authentication similar to ours but does not achieve the same performance. For instance, our system takes only 7.13 s to verify 100 signatures. In contrast, the batch verification in the cited work is limited to a maximum of 64 signatures and has verification times consistently exceeding 50 s when signatures originate from different devices. Although innovative, homomorphic encryption might not be necessary in scenarios where data are securely transmitted to trusted edge servers. Secure communication channels like SSL and TLS can provide adequate privacy without the extra computational overhead.

Discussion of Cited Studies and Their Implications

In our literature review of various IoT data authentication and privacy approaches, we identified several areas where these methods fall short, especially when compared to our approach. Key shortcomings in existing approaches include the following:

- Limited Post-Storage Verification: Some studies focus primarily on post-storage data integrity, overlooking the critical aspect of data-origin authentication.
- Over-Reliance on Edge Servers: Certain methods depend heavily on edge servers for data authentication without a clear mechanism to verify the correctness of the work carried out, raising potential security vulnerabilities.
- Permissioned Blockchain Limitations: Some frameworks utilize permissioned blockchains, which compromise the ideal of full decentralization. Additionally, when public blockchains are used, they often rely on expensive platforms without employing scaling solutions.
- Impractical Blockchain Interaction for IoT Devices: Other solutions require direct blockchain interactions using IoT devices, leading to high computational and financial costs.
- Practicality Concerns in Key Management: There are systems proposing new key management schemes that may not be practical for diverse IoT contexts. The challenge lies in adopting these new schemes as opposed to using widely used ones.
- Computational Overhead from Cryptosystems: Certain solutions, while enhancing privacy, add a significant computational workload, rendering them unsuitable for resource-constrained IoT devices.

2.2. Layered Blockchain Structure

The blockchain industry has witnessed a paradigm shift toward a layered or modular structure, notably in the public blockchain domain, such as the Ethereum ecosystem [20]. Traditional blockchain networks often employ a monolithic structure where a single chain processes and stores all transactions. While straightforward, this approach can lead to congestion and scalability issues, particularly when maintaining a high level of decentralization is a priority. In scenarios where decentralization is less critical, scalability can be enhanced by reducing the degree of decentralization. Nonetheless, as more users and transactions join the network and the emphasis on decentralization remains high, the system's capacity to process transactions efficiently is strained, leading to higher fees and slower transaction times.

Private, permissioned, or consortium networks have emerged [21], offering privacy and customization. They also enable enhanced scalability by reducing the number required for consensus. However, this shift towards more centralized control can compromise large-scale decentralization's security and resilience. Additionally, these networks often lack interoperability, isolating them from other blockchain ecosystems. This isolation results in a disconnect from public networks' innovation, liquidity, and extensive user base. This is why public networks like Ethereum have evolved to adopt a layered blockchain architecture. This includes Layer 1 (L1) networks, providing security and decentralization, and Layer 2 (L2) solutions that offer customizability and scalability by processing transactions off the main chain (L1). This layered structure can scale further into Layer 3 scaling solutions by building upon L2 solutions like rollups. This layered approach facilitates a modular and interoperable ecosystem that benefits from broader blockchain space innovations and liquidity while catering to diverse user and application needs.

Layer 2 Scaling Solutions

Multiple L2 scaling solutions exist, with rollups being the most versatile and widely used [22]. Rollups achieve scalability and lower fees by processing transactions off L1 and then consolidating batches of these transactions into a single transaction that is recorded on L1. This shared cost mechanism significantly reduces individual transaction fees. Rollups can be somewhat centralized but still inherit robust security from L1, though decentralized configurations are also possible. Additionally, rollups can be modified for a particular use case, such as allowing private data to be stored on it, essentially creating an alternative to permissioned blockchains [23].

There are generally two types of rollups: optimistic rollups and Zero-Knowledge (ZK) or validity rollups. Their primary difference lies in the type of proof they utilize. Optimistic rollups operate on the assumption that all transactions are valid unless proven otherwise, offering a reward for identifying invalid transactions. This economic incentive model, however, requires a grace period for challenges, slightly delaying transaction finality. ZK rollups, on the other hand, use complex mathematical proofs (zk-proofs) to validate transactions. They offer quicker finality and more efficient data compression, albeit with current limitations in their ability to handle complex transactions [22,24,25].

Rollups are central to our solution, offering an easier setup than permissioned ledger systems. They do not necessarily require setting up individual validators, as L1 guarantees security via proofs. This reduces complexity and enhances scalability.

2.3. Proving Systems

Proving systems are fundamental in cryptography and blockchain technology. They are used to establish the validity of statements while potentially preserving the confidentiality of underlying data. In these systems, one party, known as the prover, convinces another party, the verifier, of the truthfulness of a given statement. There are two primary types of proofs to consider:

- Validity Proofs: In these proofs, the prover demonstrates the correctness of a statement or compliance with a specific condition [26,27].
- Zero-Knowledge Proofs (ZKPs): A specialized form of validity proof, ZKPs enable a party to prove the truth of a statement to another party without revealing any additional information beyond the fact that the statement is true. This characteristic is particularly important for preserving privacy in various applications [26].

While these two categories encompass a wide array of proof types, a particularly intriguing variant within this spectrum is SNARKs, standing for "Succinct Non-Interactive Arguments of Knowledge." SNARKs are a distinctive type of cryptographic proof, marked by their two main properties: succinctness and non-interactivity. Succinctness means that the proofs are relatively compact and quick to verify, which is highly advantageous in numerous applications. Non-interactivity is a crucial feature where the prover can generate a proof without ongoing communication with the verifier.

Multiple trade-offs exist in various SNARK implementations, such as proof size, proving time and verifying time. Quantum resistance is also an important factor worth looking at [26,28].

Another important distinction is whether a SNARK is transparent or non-transparent. Non-transparent SNARKs require a process known as a trusted setup. This setup involves generating initial parameters (or keys) for proving and verifying systems. These parameters must be generated securely and trustworthy because any compromise in this process can lead to significant vulnerabilities. The trusted setup generates sensitive information which if not properly discarded, or worse, falls into malicious hands, can be used to create fraudulent proofs that appear valid. Therefore, the integrity and security of non-transparent SNARKs heavily rely on the proper execution and confidentiality of the trusted setup.

Transparent SNARKs, on the other hand, do not require a trusted setup. Instead, they use cryptographic techniques that avoid the need for a preliminary phase where sensitive parameters are generated. They are generally considered more robust and trust minimized, as they do not rely on the secrecy and integrity of a setup ceremony.

2.3.1. Snark Composition, Aggregation, and Recursion

Different proving systems offer distinct advantages and trade-offs, such as speed in generating proofs, compactness of proofs, or efficiency in verification times. SNARK composition is a powerful technique that combines these strengths to optimize all aspects of proving and verifying. In SNARK composition, multiple SNARKs are strategically merged to address their individual limitations and enhance their strengths. For example, an inner circuit with a larger size might be used to prove a computation quickly, though it might result in a larger proof size. This larger proof then becomes the input for an outer, smaller circuit designed to verify the validity of the initial statement or computation. Although the proving time for this system might be slower, its smaller circuit size and reduced proof size offer a balanced solution with relatively fast overall proving and a small final proof.

This method also enables self-composition, where a SNARK's output is recursively applied to itself. This iterative process gradually reduces verification overhead, making it particularly valuable in scenarios where verification demands more resources than proof generation [26]. SNARK composition enables significant applications like Incremental Computations and Proof Aggregation. Incremental Computations use SNARKs to verify the correctness of each step in a longer computational process, ensuring integrity throughout. Proof Aggregation involves combining multiple proofs into a single, more manageable proof, which is especially useful in scenarios requiring collective verification of numerous individual proofs.

Figure 1 illustrates how proof aggregation might be applied to batch signature verification. The naive approach would be to accumulate a large number of signatures before verifying them and create a single comprehensive proof of verification. With proof aggregation, proofs can be generated more promptly, and all the created proofs would then be aggregated, essentially creating a "proof of proofs".

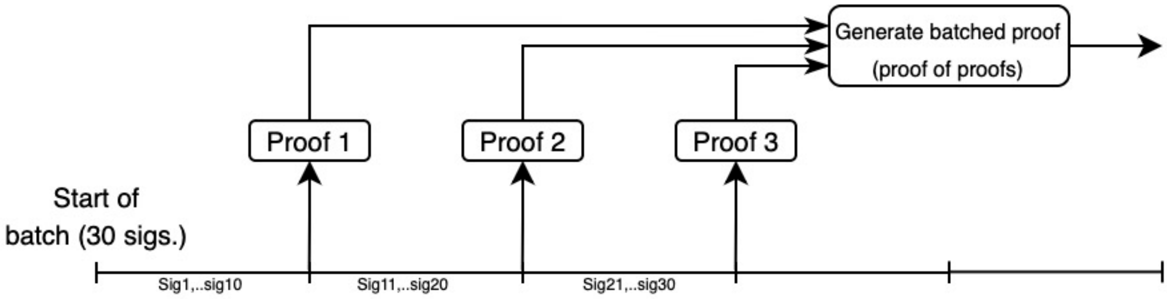

Figure 1. Proof aggregation.

Recursive proofs, on the other hand, involve iteratively applying a function F (such as signature verification) to a computation with private inputs, where each step given private inputs (ω) generates a proof (π) confirming the accuracy of the current computation relative to the previous state (s_{n-1}). Known as Incrementally Verifiable Computation (IVC), this method reduces memory usage by generating smaller, individual proofs at each step instead of a single, extensive proof or aggregating multiple proofs into one (Figure 2).

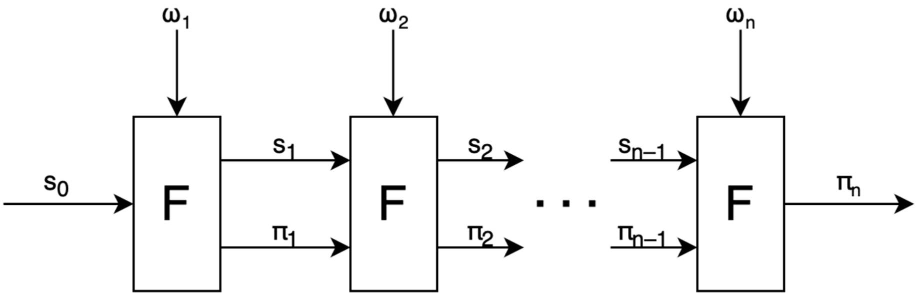

Figure 2. Incremental verifiable computation.

However, this approach also presents challenges. Since each new proof must validate the correctness of the previous proof, the process becomes computationally intensive. Despite the approach's power and flexibility, creating and verifying a proof at every step increases the overall computational burden.

The Nova prover [29,30] introduces an innovative solution to the challenges of recursive proofs. Nova's technique involves composing proofs throughout the computation but only finalizing them at the end. This strategy substantially reduces computational overhead, as it incurs the cost of proof generation just once.

Nova employs a process called Statement Folding to merge multiple instances into a single instance for verification. In this method, the prover P constructs a proof for a circuit C without needing to run the entire verification algorithm V(Vk, X, π), which typically involves the verifier key Vk (a part of the public parameters), the statement X, and the proof π, with the output being true or false. Nova simplifies circuit C by removing most verification checks, meaning that the prover only needs to construct a proof for a few selected checks, greatly simplifying the proof construction process. While Nova benefits from needing to verify only the last compressed SNARK, it also must ensure the correct execution of the folding process. This verification is integrated into the computed and proved function, typically referred to as the accumulator, which essentially accumulates and verifies that the folding was indeed completed correctly.

2.3.2. Signature Aggregation

Since our solution focuses on removing interactions between IoT devices and the blockchain, there is a need for an alternative approach to signature verification. Traditionally, this was conducted by verifying the IoT signature directly on-chain and storing the actual signed data/messages off-chain. This approach was employed to lower data storage costs on-chain, but signatures were not eliminated. They still needed to be recorded on-chain and verified via a smart contract.

Signature aggregation allows us to remove even the signature verification from the chain. Taking care of the signature verification and data storage off-chain drastically lowers costs and paves a new way for optimization that does not require as much intervention with blockchains. It essentially moves the bottleneck from the blockchain to an off-chain system that can be further scaled using traditional means, even if that means adding extra hardware. This flexibility is not possible on current blockchains, where we are limited by the performance of any single chain.

In essence, signature aggregation is a method where multiple signatures from different sources (in this case, IoT devices) are combined into a single, compact signature. This process significantly reduces the amount of data required to verify the authenticity and integrity of messages from multiple devices. A common approach in blockchains, for example in Ethereum, is BLS signature aggregation [31], which allows the reduced signature size to be verified against the aggregated public key, created by combining all the public keys from the IoT devices.

While this approach is interesting, the fact that the majority of IoT devices use elliptic curve cryptography (ECC), specifically ECDSA, necessitates a different approach to aggregation. This is where SNARK aggregation comes into play, allowing us to replace many signatures with a single SNARK proof that validates the authenticity of all signatures.

As mentioned in Section 2.3.1, the Nova prover demonstrated how we could recursively prove any computation. In this case, the function F we are recursing would be the signature-verification algorithm. However, proving any computation is not trivial, and it is computationally more intensive than running the algorithm on its own. Modifying the signature verification to be more efficient, therefore, provides considerable performance gains. The efficient ECDSA verification method from Personae Labs [32] achieves just this. The performance benefits stem from executing certain computations outside of the proving phase. In essence, it is an optimized process that restructures the ECDSA signature verification so that parts of it are computed off the SNARK circuit. This method works by rewriting the ECDSA signature-verification equation to isolate elements that can be computed outside the SNARK, thus reducing the number of operations within the SNARK itself.

3. Framework Design

Our proposed framework enables a performant, secure, and privacy-focused IoT data authentication and storage method, utilizing permissionless blockchains and Zero-Knowledge proving systems. This framework achieves its goals by implementing a secure data validation process that minimizes the computational burden on IoT devices, distinguishing our solution with its performance-oriented design and adaptable architecture. It is set up to scale effectively, capable of accommodating an increasing number of IoT devices and data volumes. This scalability is achieved either by adding new L2 rollups or expanding the network of edge aggregating servers, ensuring the system adapts and grows without compromising performance, security, or cost-efficiency.

The solution comprises two main components:

- Off-chain components include IoT devices and edge aggregating servers. The aggregators serve as the first point of data aggregation and authentication, collecting and validating data from IoT devices, thus preparing it for on-chain integration. This process significantly reduces the computational load on individual IoT devices and decreases the need for direct blockchain interaction.
- On-chain components feature a layered blockchain structure and smart contracts built atop each L2 network. Ethereum is the foundational Layer 1 (L1), ensuring secure and reliable operations, while Layer 2 networks, such as Optimism Rollups, provide scalability and efficiency. Smart contracts facilitate IoT and edge-server registration and data storage.

3.1. Off-Chain

In this subsection we will delve into the off-chain part of our solution in more detail. Starting with IoT devices and data authentication via signatures and later using a proving system on the edge servers to aggregate signatures. In terms of IoT devices, we assume that the devices themselves are secure from a hardware perspective and/or on the software level, allowing the safe storage and use of private keys and data generation. Similarly, for edge aggregating servers, we also assume security at both the hardware and software levels, which is crucial for the effective operation of the proving system and the reliable aggregation of signatures.

3.1.1. IoT Devices

To enable data authentication, all IoT devices need to do the following two tasks:

- Initial on-chain registration (only once);
- Data authentication via digital signatures.

Blockchains offer a solution that prevents data manipulation because of the immutable nature of blockchains. This is why every IoT device has to authenticate itself on the ledger, resulting in the creation and storage of its certificate, which includes its public key and other possible identifying details (e.g., what type of IoT device it is and what data does it produce). These certificates (public keys) are used to make sure that the IoT devices indeed generated the data.

Some IoT devices might be resource constrained, so expecting all IoT devices to communicate directly with a chain for anything else besides initial registration is not viable, since it comes with additional financial costs (paying for transaction fees).

To alleviate IoTs from this job, we introduce edge servers that aggregate IoT signatures into one single signature/proof. The signatures are thus essential for our solution, where we expect IoT devices to sign their messages before sending them to the edge aggregating server. Using digital signatures on the IoT side with verification on the edge aggregator offers faster data authentication. However, it increases the computational load on the IoT device, which may not be ideal for resource-constrained devices, such as those running on battery power. For these resource-constrained devices, Hash-based message authentication code (HMAC) is available, which reduces the computational burden on IoT devices by sacrificing the speed of data authentication.

3.1.2. Edge Aggregating Servers

Our solution employs cryptographic proofs (SNARKs) and the algorithm used for ECDSA signature verification within the Nova prover. This enables a coprocessor of sorts on top of the blockchain. These coprocessors, or edge aggregators, allow for the offloading of computation and storage burdens from the blockchain to off-chain systems. This approach helps to decouple unnecessary functions from the ledger, moving bottlenecks to off-chain systems where they are more manageable. The ECDSA verification algorithm checks the validity of signatures, and the Nova prover confirms that the algorithm has been executed correctly and on valid inputs, including signatures, messages, and public keys. The resulting proof serves as an aggregate verification of the signatures' authenticity. It is important to note that while this method increases overheads on off-chain systems, it is more feasible than handling these tasks on the blockchain, where resources are more limited.

The operational steps each edge aggregator has to make are as follows:

- Initial on-chain registration (only once).
- Signature pre-processing (conform with efficient ECDSA).
- Building a Merkle tree for each batch (includes public keys, signatures, messages, and batch identifiers) ensures all relevant data are accounted for in a verifiable structure.
- The edge aggregator signs the Merkle root along with the batch number and uses this as the public input to the prover. This ties the proof generation directly to the specific dataset represented by the Merkle tree.
- Writing each batch's Merkle root, proof hash, and batch identifier on-chain. It enhances transparency and provides an immutable record that can be independently verified.

Aggregating Signatures

All edge aggregators must register on-chain the same way IoT devices do. This is because they are responsible for aggregating the signatures and writing on-chain, so we need a way to verify that the right aggregator responsible for a set of IoT devices has indeed completed the work and published the relevant data (for authentication) on-chain.

At the core of our system is the process of data collection from IoT devices using the edge aggregators. Each IoT device collects data, referred to as 'messages', and signs these messages with its private key to authenticate their origin. All these data are then sent to an edge aggregator where some pre-processing is completed using the signatures and actual message to conform with the efficient ECDSA verification method explained in Section 2.3.2. These signed messages are then batched for processing, with each batch being assigned a unique identifier to distinguish it from others. This unique identifier plays a crucial role in the later stages of verification.

For each batch of data, a Merkle tree is constructed. The leaves of this Merkle tree consist of a hash that incorporates the message, its signature, the corresponding public key, and the batch identifier. The construction of the Merkle tree is a critical step as it enables the efficient and secure verification of the data inclusion in the batch. The Merkle tree root and the current batch number are used as the message the aggregator has generated, so it has to sign it. Pre-processing is completed for this message and signature as well, so it can be used as the public input in the proving system (Nova). All other IoT data will be the private inputs. This way, we find a ZK proof that gives no information about the IoT data. The public input, which is the signed Merkle tree root plus the batch number, allows us to prove that a certain IoT data packet was part of the authentication process via signature aggregation.

The size of the aggregated signature, or more exactly, the proof of aggregation, also called the recursive proof, is quite large, so an additional proving is completed to compress the recursive proof. This compressed SNARK proof comes with the Nova prover and is based on Spartan. Essentially, it creates a smaller proof that proves that we have a proof that satisfies the initial statement, in this case, that the aggregation has been completed correctly. The proof is intrinsically linked to the batch identifier, ensuring a clear and verifiable connection between the proof and the specific batch of data.

To enhance the integrity and provide a timestamp for our system, the aggregator records crucial information on a blockchain. This includes the hash of the aggregated signature (compressed cryptographic proof), the Merkle root of the current batch, and the batch identifier. Using a blockchain here is pivotal, as it offers an immutable ledger, ensuring that the data cannot be altered retrospectively once recorded.

To see how the solution would work, imagine a transaction occurs and a buyer purchases specific data from an IoT device. They are provided not just with the message but also its signature, the corresponding public key, the batch identifier, and crucially, a Merkle proof. This Merkle proof enables the buyer to independently verify that the specific message was indeed part of the batch linked to the recorded proof. The buyer can verify this by checking the Merkle proof against the publicly recorded Merkle root and ensuring that the batch identifier aligns with the blockchain record. This verification process confirms that the message was part of the specific batch for which the proof was generated, thus assuring the buyer of the authenticity and integrity of the data. Verification that the proof is correct can also be completed by providing the buyer with the proof (along with necessary verification tools) and them verifying it themself.

Security considerations of this approach are discussed in Section 5.

Data Storage

The edge aggregating servers also take care of storage (IoT data like sensor readings) by storing them on the aggregator itself, in a cloud database, or completely on-chain. The image below depicts the whole system where the aggregator writes the final proof/aggregate on-chain. The idea here is that once a setup has been created, the aggregator keeps producing new proofs for batches of signatures and then writes each proof and IoT data on-chain. Smart contracts are deployed to store the data and verify the proof. While this approach would be the ideal solution, it is not yet possible as of this paper, with the reason being that the proof and whole data are still too big. Consequently, high gas costs are accrued for writing all of these data. The complexity of creating a verifier smart contract is also an issue. We believe that the exploration of full on-chain storage should still be considered and explored further since rollups on L2 can also be built to store data on cheaper systems than on L1. Perhaps the use of Layer 3 or alternative blockchain storage solutions like EigenLayer, Celestia, Filecoin, and Arweave could also be considered (Figure 3).

A better approach would be to store all of the IoT data off-chain and only include hashes of the proofs on-chain, working as a timestamp of when the proof was generated and preventing the aggregators from tampering with the proofs. From there, anyone can verify the proofs on their own, which is an acceptable solution for many non-financial privacy applications. The image below depicts such a system (Figure 4).

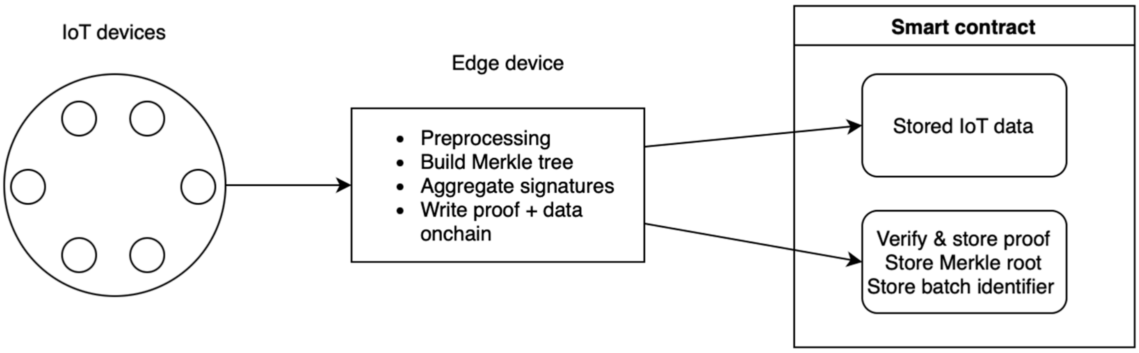

Figure 3. Fully on-chain storage.

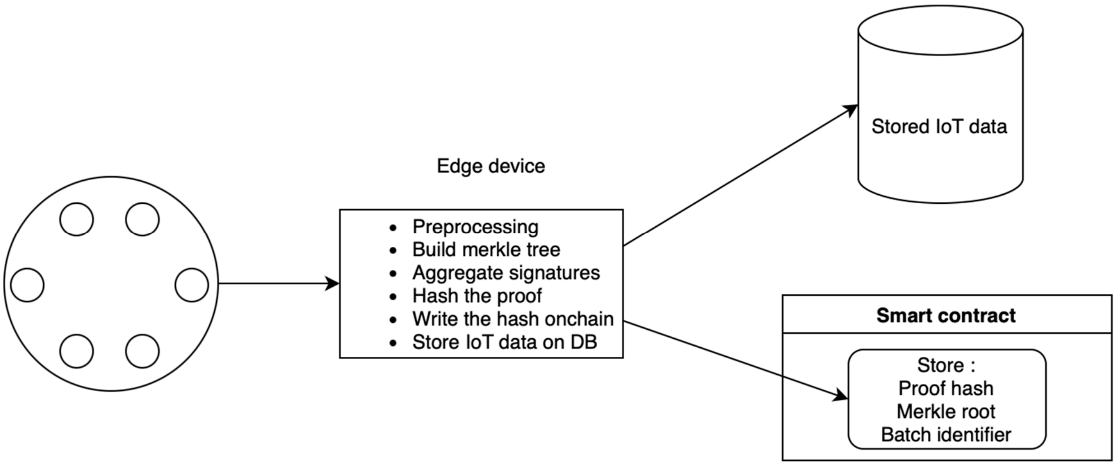

Figure 4. Divided storage (on-chain + off-chain database).

3.2. Proving System Implementation

The implementation and testing of our aggregator started with choosing a proving system that would allow us to skip the trusted setup that comes with some SNARKs. We first utilized the RISC Zero zkVM, a high-performance tool for proving the correct execution of arbitrary code written in Rust. This tool lowers the threshold to achieve a working provable program because we just have to write Rust code and the zkVM proves it. Applications written in Risc0 are structured into two parts: the host code and the guest code. The guest program is the code that we want to run; in our case, it is the ECDSA verification algorithm, and the host code is the code that proves that the guest code was executed correctly. The idea here was to create one proof of verification per signature and lastly create an aggregated proof of proofs with their recursive solution that would compress many Risc0 receipts (proofs) into a single receipt and lastly compress the last receipt via their STARK-to-SNARK circuit that translated a STARK proof into a SNARK proof. Unfortunately, their recursion circuit was not yet available or documented at the time of writing and neither was their compression circuit. Nonetheless, the proving time for one ECDSA signature verification was an eye opener since the proving speed was unsatisfactory for large IoT data authentication. Despite the ease of use of Risc0, the lack of performance led us to Nova, a performant recursive prover.

In Nova, we have to write our own circuits. We re-used the Efficient ECDSA verification circuit written by Personae Labs with some modifications. The circuit was written in Circom, but Nova natively supports the Bellpeperson library for Rust, so we utilized an open source library, Nova-scotia, as the middleware between Nova and Circom. Since we are verifying ECDSA (secp256k1) signatures, we utilized secp/secq curve cycles, which Nova already support. The proving side works over the secq curve and the verifying works over secp, so when compiling with the Circom compiler, we had to specify the prime to be used—prime secq256k1 for the circuit generation.

The outputs are the circuit in R1CS and the witness generator in wasm. We could then write the implementation for Nova in Rust. We read the pre-processed signatures, generated the public parameters using the R1CS file, and lastly, created and verify the proofs.

There are two proving stages: the first is the recursive one that proves the validity of all signatures (signature aggregation), and the second for compressing the recursive proof. The recursive proving takes in the public input s_0 and the private inputs ω (witnesses), with F being the signature aggregation. The public input is the Edge serve message, signature, and public key, and the private inputs are all of the IoT messages, signatures, and public keys. Both the public and private inputs are pre-processed to conform with the efficient format (Figure 5).

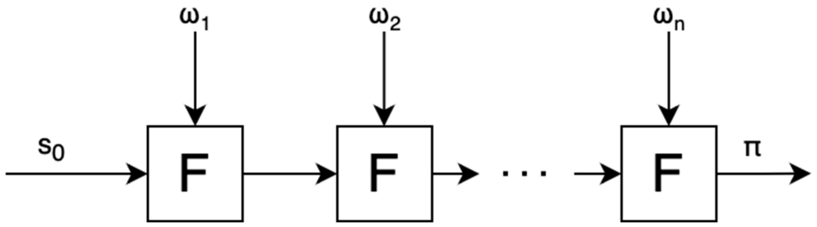

Figure 5. Nova recursive prover.

Alongside the actual inputs, the recursive prover also takes the R1CS and witness generator files, as well as the public parameters generated with the R1CS file.

For each aggregation step, the prover will take 10 signatures and then continue to the next step with the next 10 signatures. At the very end, we receive the recursive proof, which is quite large, which is why we further compress this proof with the Spartan prover incorporated with Nova.

3.3. On-Chain

3.3.1. Layer 2 Rollups

We have set up a Layer 2 (L2) rollup for our proposed solution, since we believe it presents a superior approach for real-world applications. This preference is due to rollups' ability to remain centralized and thus gain scalability while still ensuring security by posting proofs and compressed block data on the primary Layer 1 (L1). Initially, we set up our rollup to settle on Ethereum's test network Holesky to assess usability and functionality.

However, for the actual testing and benchmarking, where costs of interacting with the rollup were recorded, we employed the Optimism Sepolia testnet. This choice was made to ensure more realistic blockchain load conditions, which could not be replicated on a laptop with minimal traffic. The Sepolia testnet provided a more accurate environment for testing, with loads similar to real-world scenarios.

This layered structure, where we can spin up new rollups, facilitates the creation of new modules that can achieve better performance and privacy (with some modifications) while relying on L1 for settlement and security. As one rollup becomes congested, we can simply spin up new rollups, all settling on Ethereum. This approach is also intriguing because we do not silo ourselves from the broader ecosystem, which includes an abundance of open-source developments, liquidity, and user bases. Furthermore, this model provides a standardized means of interoperability with other rollups. For instance, two companies could create their L2s with tailored adjustments and exchange data and resources within the shared ecosystem via bridges or Ethereum itself.

3.3.2. Smart Contract

While the layered blockchain structure allows for increased modularity and scalability by spinning new rollups, additional modularity should be achieved on the smart contract side. We implement this by utilizing the MUD framework [33]. It allows for our proposed system to be scalable, handling potentially vast numbers of IoT devices that are aggregated using multiple edge aggregators.

The MUD framework comprises two primary components: Store and World. Store serves as an alternative to Solidity's storage engine, offering a data model similar to a relational database or key-value store. This design allows for automatic indexing via event emissions on each storage operation and packs data more compactly than Solidity's storage engine. Moreover, it enables on-chain reading of external contract storage without depending on existing view functions.

In our implementation, each piece of IoT data (hashes, proofs, and/or the data itself) is stored as a record in a table within Store. Each record is uniquely identified using a combination of a ResourceId (tableId) and a composite key (bytes32[] keyTuple). The ValueSchema of each table defines the data types stored, similar to column types in a database table. The table schema used in our implementation is outlined in Table 1, displaying the schemas for device registration and aggregator data entry. As the network of IoT devices grows, requiring more edge aggregators, our system can expand by creating multiple namespaces with these tables. Access control for each namespace ensures that only specific edge aggregators can write batches.

Table 1. Schema for our data tables.

Device Registration	**Aggregator Data Entry (Batch Info)**

keySchema: {
owner: "address",
},
valueSchema: {
isIotDevice: "bool",
},
	keySchema: {
	owner: "address",
	},
	valueSchema: {
	batchNumber: "uint32",
	proofHash: "string",
	merkleRoot: "string",
	},

Store also automatically generates a library for each table, providing getter and setter functions. Furthermore, Store allows for runtime schema definition, enabling the registration of new tables with new schemas after deployment. This capability is crucial for advanced applications such as the World protocol, which we utilize in our IoT data authentication framework.

The World component of MUD provides the logic and access control layer on top of Store's storage capabilities. It acts as a central entry point for calls, performing access control checks and routing authorized requests to the appropriate System. Systems are stateless contracts interacting with data in Store. They can read from all tables and modify data in accessible tables, typically within their namespace, which can be understood as containers for tables and systems. Even access control is based on namespaces.

Our experimentation involved several steps. Initially, we defined the table schemas. This was followed by the implementation of system contracts, utilizing the libraries generated for each table. We crafted a single system contract containing functions to modify the data in both tables. Subsequently, we deployed the smart contracts using the MUD framework. After deployment, our focus shifted to logging the gas costs associated with device registration and the entry of edge aggregator data for specific batches. The entire process was straightforward, effectively highlighting the user-friendly nature of the MUD framework.

Figure 6 depicts the complete solution architecture. The process begins with multiple IoT devices collecting data and generating corresponding digital signatures. These data, along with the signatures, are then transmitted to the edge aggregating servers.

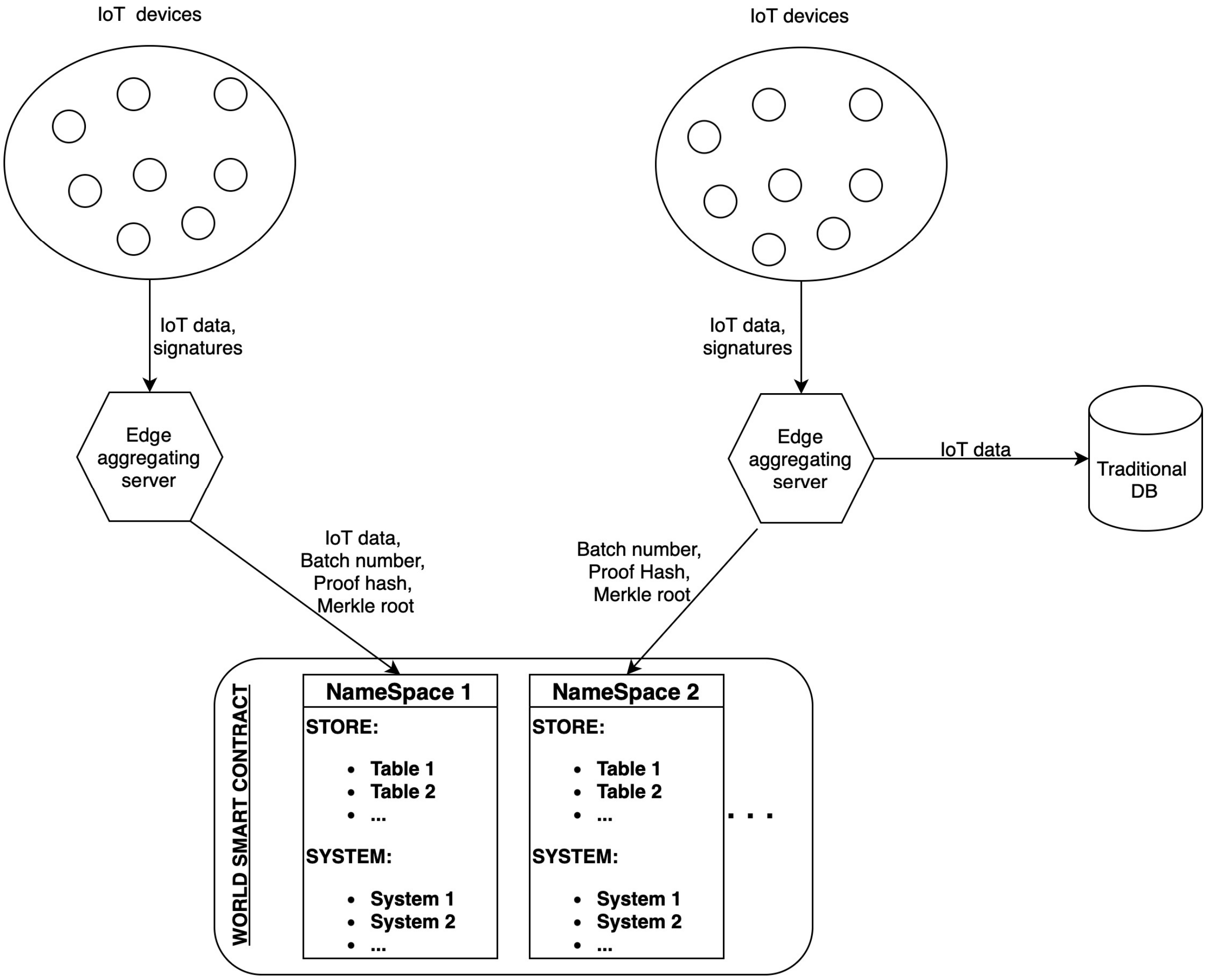

Figure 6. Overview of the whole solution (smart contract + off-chain aggregation).

On the left, we have illustrated how a fully on-chain storage approach would look like, but it is important to note that we have not tested this approach. We have instead defined alternative research approaches in the Conclusion section (Section 6) that focus on a fully on-chain approach. The edge aggregator receives data and signatures from its connected IoT devices. It performs signature aggregation to create a compact batch that includes aggregated IoT data, the batch number, proof hash, and Merkle root. This batch is then written to NameSpace 1 on the blockchain for immutable storage, ensuring integrity and verifiability of the data. On the right, another edge aggregator similarly receives data and signatures but opts for a different data handling approach. After aggregation, the IoT data are forwarded to a traditional database for storage, while only the batch number, proof hash, and Merkle root are written to NameSpace 2 on the blockchain. This method segregates the detailed IoT data from the blockchain, leveraging traditional databases for storage, and uses the blockchain primarily for verification purposes.

Below the aggregators, the diagram shows the World Contract, which acts as the backbone of the system's on-chain component. It houses NameSpace 1 and NameSpace 2, each containing tables for storing data and systems for executing logic.

4. Results

Our solution is structured into two main components, resulting in our findings being split into two distinct sections: off-chain and on-chain. Our experimental setup for the off-chain (edge aggregating server) component was carried out using a MacBook M1 Pro with 16 GB of memory, which facilitated the creation of signatures, their verification, and the generation of proofs. Our framework is designed to be compatible with any IoT device capable of securely creating signatures. This includes devices that can safely store and use private keys. For the on-chain component, we utilized the public Optimism Sepolia testnet, rather than our own rollup, to simulate more realistic blockchain load conditions and provide a more accurate assessment of on-chain costs.

The off-chain portion elaborates on the signature aggregator's proving times and proof sizes. This aggregator processes IoT messages and signatures, compiling them into batches, generating succinct proof of verification/aggregation. In this context, we have explored two proving systems, both employing a transparent approach that eliminates the need for a trusted setup. We present a comparison of the outcomes achieved by these two systems. On the other hand, the on-chain section delves into the storage costs and the registration process for devices, encompassing both edge aggregators and IoT devices. Our tests on the on-chain component focused solely on recording the essential data needed to authenticate an entire batch after its proof. While direct on-chain storage of all IoT data has not been evaluated, the costs can be inferred from the expenses incurred in storing only the necessary batch data.

4.1. Proving Times and Proof Sizes

Figure 7 illustrates the comparative proving times for the Nova and Risc0 systems as the number of signatures required for aggregation increases. For Nova, the proving times represent the total of two components: the time required to generate the recursive SNARK and the time to produce a smaller, compressed SNARK derived from the recursive one.

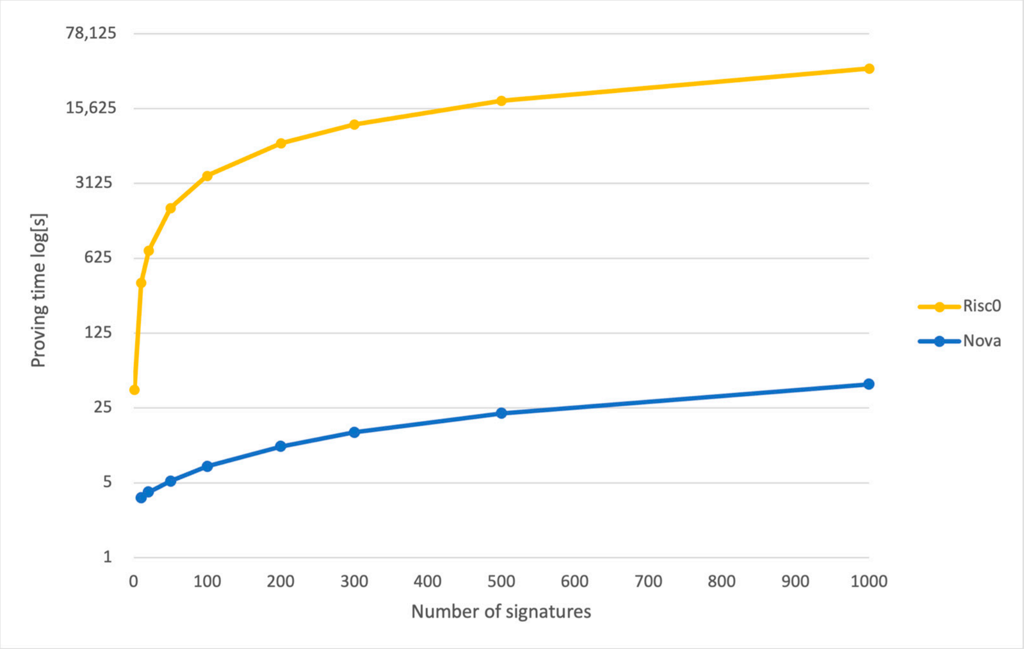

Figure 7. Proving time for Nova (recursive + compressed) and Risc0.

Figure 8 presents the proving times for the recursive and compressed SNARKs in the Nova prover. We can see that the time for the compressed SNARK remains constant, while the recursive one rises linearly with the increase in the number of signatures. This is because the compressed one only proves the correctness of the recursive SNARK regardless of the number of signatures, while the recursive one has to aggregate more signatures.

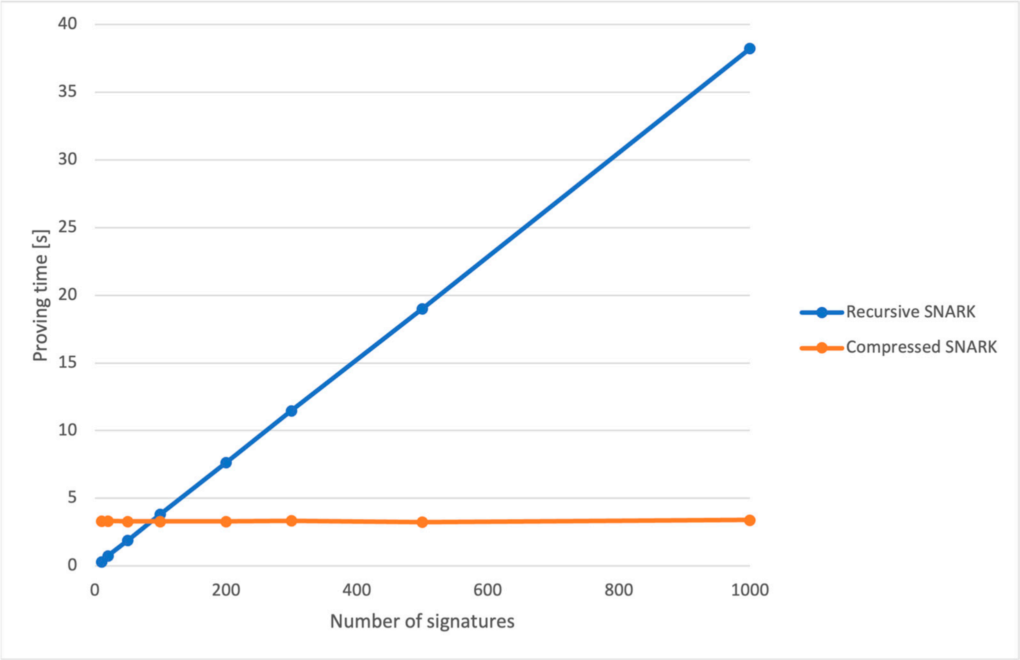

Figure 8. Nova proving time for the recursive and the compressed SNARK.

The recursive proof in Nova is approximately 8.7 MB, while the compressed proof is just 29, in contrast to the 1.3 MB size of the Risc0 proof. However, it is important to note that the Risc0 proof is not compressed, so for our case, aggregating multiple proofs and then compressing the last proof would be necessary.

While proving time is important on the edge aggregating server side, verification times are crucial for clients who need to verify that the aggregation and, thus, data authentication have been correctly performed. Figure 9 depicts the verification time per signature for both the Risc0 and the Nova provers. We have combined the verification times for the Nova prover for the recursive and compressed SNARKs. It can be observed that the verification times per signature remain constant for the Risc0 prover, which is expected as a proof is created for each signature verification.

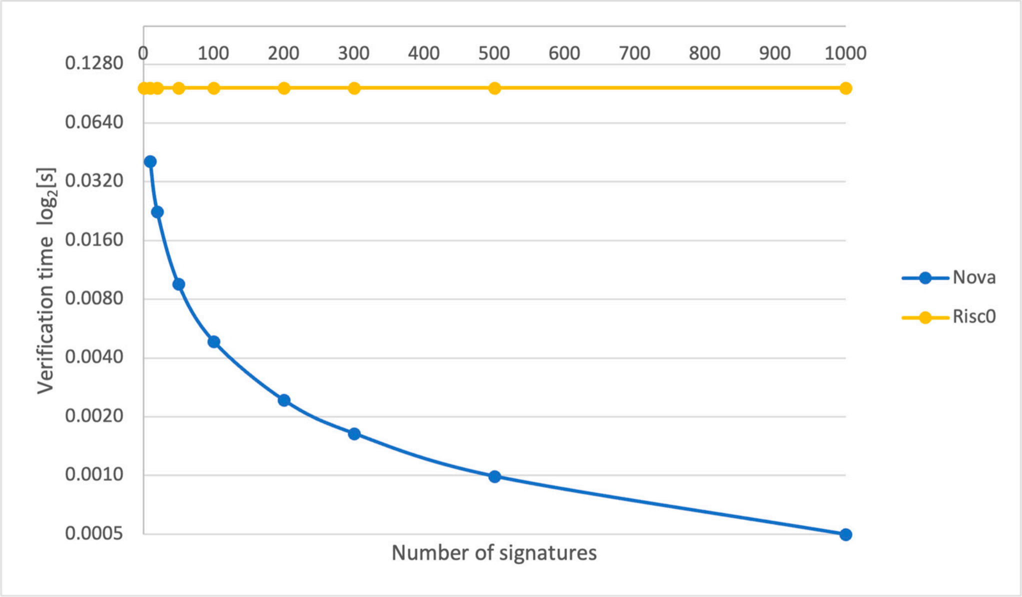

Figure 9. Verification time per signature for Nova (recursive + compressed) and Risc0.

Analyzing the proving times per individual signature is equally important. That is why Figure 10 illustrates the proving times for both the Nova and Risc0 proving systems on a per-signature basis. It shows that the proving times for Nova decrease as the number of signatures in a batch increases, indicating improved efficiency at scale. In contrast, the proving times for Risc0 remain constant, which is consistent with its overall proving time trend.

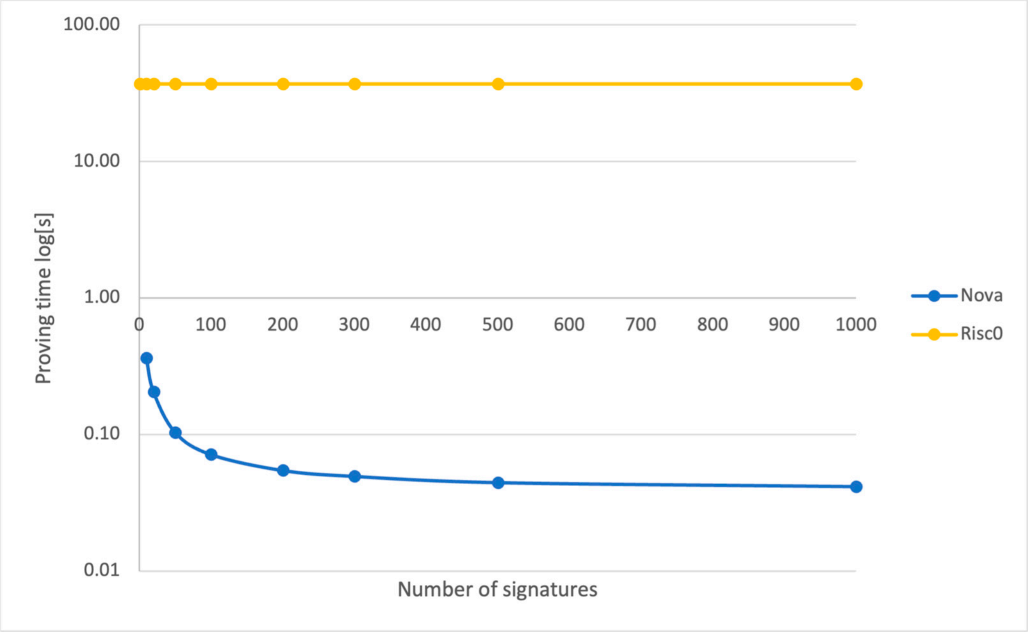

Figure 10. Proving time per signature for Nova and Risc0.

4.2. On-Chain Storing Costs

For our on-chain test, we assessed the costs associated with registration and the cost of writing authenticating data for each batch, which includes the batch number, aggregating proof hash, and the Merkle root. We present the costs for Layer 1 (L1) and Layer 2 (L2) networks, demonstrating how using L2 can significantly reduce costs.

Two factors contribute to gas consumption within L2 networks, specifically in our case, with a rollup based on the Optimism stack. The first is the portion of gas used for writing our transaction to L1, and the second is the gas used for executing the transaction on L1. The gas prices for these two components differ.

Equations (1) and (2) depict how the transaction fee is calculated based on the gas used for the transaction and the current gas price for both networks. This calculation includes the base gas price, the priority gas price, and a scalar representing a dynamic overhead cost, which, at the time of writing, was set to 0.684. To determine the actual transaction fee, both fees from L1 and L2 must be summed.

$$transaction\ fee\ on\ L2 = gas\ used \cdot (base\ price + priority\ price) \tag{1}$$

$$transaction\ fee\ on\ L1 = gas\ used \cdot (base\ price) \cdot scalar \tag{2}$$

The prices for the gas units and ETH utilized in the calculations are detailed in Table 2, which reflects the average market rates for the Ethereum and Optimism networks at the time of writing, with 1 ETH being equivalent to 10^9 gwei units.

Tables 3 and 4 detail the costs of a single registration and data writing. The data are presented in terms of gwei units and US dollars for two distinct scenarios: one using the Ethereum mainnet and the other employing a rollup based on Optimism.

Figure 11 depicts the difference in accumulated costs for the scenario where we write the authenticating data for a batch of signatures. It shows how the costs of utilizing L1 instead of L2 would rise dramatically as the amount of data written for batches rises or accumulates over time.

Table 2. Prices used in cost calculations.

Layer 1		Layer 2		
Base Price [gwei/gas]	**Priority Price [gwei/gas]**	**Base Price [gwei/gas]**	**Priority Price [gwei/gas]**	**ETH Price [USD/ETH]**
28	0.1	0.00345	0.02	2150

Table 3. Gas usage and cost for device registration in the case of using L1 or L2.

	Layer 1 Only	**Layered (L1 + L2)**	
	Gas Used	**Gas Used L1**	**Gas Used L2**
Registration	74,653	2188	74,653
Cost in gwei and USD			
Cost [gwei]	2,097,749.3	41,904.576	1750.612
Cost [USD]	4.51	0.090	0.003
Cost total [USD]	4.51	0.093	

Table 4. Gas usage and cost for batch data writing in the case of using L1 or L2.

	Layer 1 Only	**Layered (L1 + L2)**	
	Gas Used	**Gas Used L1**	**Gas Used L2**
Writing data	198,979	4820	198,979
Cost in gwei and USD			
Cost [gwei]	5,591,309.9	92,312.64	4666.057
Cost [USD]	12.02	0.198	0.010
Cost total [USD]	12.02	0.208	

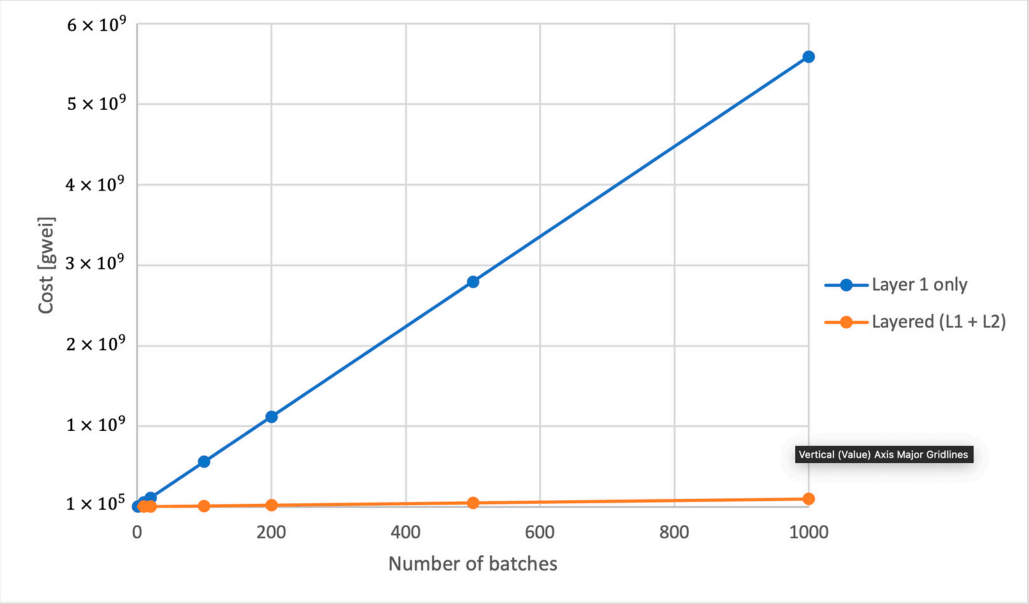

Figure 11. Cost comparison between using Layer 1 only and Layered (L1 + L2) for writing batch authenticating data.

5. Discussion

IoT data authentication, traditionally centralized, has been revolutionized using blockchain technology, offering immutability and enhanced security. However, while secure, direct interaction with public ledgers is not economically scalable due to transaction costs. Permissioned blockchains improve scalability and privacy, but at the cost of true decentralization, reintroducing a degree of trust and thus lower security.

Our approach tackles this problem by preserving high security while still making the system highly scalable. The solution consists of an off-chain edge aggregating server that manages IoT devices and handles authentication via signature aggregation, coupled with a layered blockchain structure. The first layer includes the Ethereum network, providing high decentralization and security, while the second layer comprises rollups with higher performance and cheaper fees. These rollups are used for device registration—encompassing IoT devices and Edge aggregation servers—and for storing authenticating data for each batch of signatures the edge aggregator processes. These data include the proof hash produced by the proving system inside the edge aggregator, the Merkle tree root, and the current batch number.

5.1. Off-Chain Results

We utilized the Nova and Risc0 provers, with Nova being used to aggregate batches of signatures and produce a proof at the end of aggregation and Risc0 to create a proof of signature verification for one signature only. Figure 7 provides a logarithmic comparison of the proving times for Nova and Risc0 as the number of signatures increases. Utilizing a log scale is crucial to effectively visualize Risc0's performance in conjunction with that of Nova, given that Risc0's proving times escalate much faster with an increasing number of signatures. For example, when the count of signatures rises from 10 to 20, Risc0's proving time nearly doubles from 369.7 s to 739.4 s, while Nova's proving time grows from 3.62 s to only 5.187 s.

The substantial performance gap arises because generating a proof for each signature verification is computationally intensive. Risc0 operates as a non-recursive prover, creating an individual proof for each signature verification. In contrast, Nova functions as a high-performance recursive prover, verifying batches of signatures and producing a proof only after all verifications are complete. This difference in approach underscores the scalability and efficiency of Nova, particularly in scenarios with a large number of signatures.

This gap becomes more pronounced when comparing both systems for proving times and verifying times per signature (Figures 10 and 11), where the Nova prover demonstrates increased efficiency with more signatures, unlike the constant times of the Risc0 prover.

However, the Nova prover's efficiency comes with a complexity: it has two contributors to proving time. The first is the time it takes to create a recursive proof, and the second is the time required to compress it. These two contributions are there because the proofs created from recursion were too large, so another step of proving the recursive proof was used, creating a final compressed proof which is much smaller in size. For instance, the recursive proof in Nova is approximately 8.7 Mb, while the compressed proof is just 29 Kb, in contrast to the 1.3 Mb size of the Risc0 proof. In Figure 8, we noticed that the proving time for the recursive side grew linearly as the number of signatures being aggregated rose, while the time for compressing the proof stays the same. This makes sense, since the compressing part always takes in one proof regardless of the number of signatures.

It is important to note that the proofs generated with Risc0 could be combined into a single succinct proof. However, we chose not to pursue this due to incomplete Risc0 documentation on this subject and the lengthy proving time for one signature verification. The Nova prover was thus chosen as it provided excellent performance, with proving and verification times per signature decreasing as the number of signatures in a batch being processed grew.

5.2. On-Chain Results

For the on-chain results, we measured the costs associated with registering devices and writing the authenticating data from the edge aggregator for one batch, irrespective of the number of signatures in a batch, as the proof hash and Merkle root sizes remain constant. The results demonstrate the cost advantage of using Layer 2 rollups over Ethereum Layer 1. Specifically, the costs of using Layer 1 only are almost 48 times larger than using Layer 2 for device registration and 57 times larger for authenticating data writing. Thus, employing a Layer 2 rollup significantly reduces costs compared to using the Ethereum mainnet. This cost reduction is in addition to the decreased costs for IoT devices, which, in our approach, do not need to interact with a ledger. Yet, the proving system and the Merkle tree root guarantees that the data are authentic, originating from registered IoT devices, and have not been tampered with.

5.3. Security Discussion

With our approach leveraging cryptographic proofs (SNARKs) and Merkle trees, we have established a foundation for secure and privacy-focused IoT data authentication and storage. This method inherently enhances security: any tampering with the input data or the prover would result in the generation of invalid proofs, thereby signaling potential security breaches and ensuring that only valid data are processed and recorded.

Nevertheless, for a system to be production-ready, other security considerations must be taken into account. While our research assumes the security of IoT devices and edge servers at both the hardware and software levels, the practical implementation of such systems in real-world scenarios may present additional challenges. Ensuring the physical and digital integrity of IoT devices is crucial, as they are often deployed in uncontrolled environments and could be prone to tampering or unauthorized access. Similarly, edge servers, while assumed to be secure in our framework, would require rigorous security protocols to prevent data breaches, unauthorized access, and ensure data privacy.

Another critical aspect is the secure communication between IoT devices and edge aggregating servers. It is essential to safeguard data transmission to prevent interception, eavesdropping, or manipulation. Implementing robust encryption protocols for data in transit, such as TLS/SSL, along with mutual authentication, integrity checks, and access controls, can provide the necessary multi-layered security.

Further research should explore these areas in depth, focusing on robust methods to secure IoT devices and edge servers against a wide range of threats. One potential avenue is the integration of automated security checks that could continuously monitor for signs of tampering or unauthorized access, both in IoT devices and at the server level. Such mechanisms would be particularly effective given the security guarantees our approach allows. The use of SNARKs and Merkle trees not only enhances data integrity but also offers a solid foundation for developing additional security features.

Furthermore, while the approach is discussed in Section 3.1.2 using Merkle trees enhances the verifiability of the data, there is still an element of trust in the server. The server's role in constructing the Merkle tree and generating the public input for the prover must also come with the guarantee that the private inputs going into the prover are the same data being used to build the Merkle tree. As of this paper, this issue has not yet been tackled, but some approaches we foresee could be as follows:

1. Modifying the Prover.
    a. Modify the prover to take public keys as inputs and signatures and messages as private inputs. This could provide a more direct link between the data sources (IoT devices) and the proof.
    b. Modify the prover to construct the Merkle tree as part of the proof generation process. This would tightly couple the data verification with the proof itself, providing strong assurance that the same data are used in both the Merkle tree and the proof, without the need to disclose any data for each batch.

2. Proving that the process of building a Merkle tree used the same data as the process of aggregating signatures. This could be completed by utilizing yet another proving system or for example, Intel's trusted execution environment (SGX).

6. Conclusions

In contrast to existing IoT data authentication methods described in Section 2.1, our solution offers a unique blend of efficiency, scalability, and privacy. Unlike approaches that primarily rely on post-storage data integrity checks or over-dependence on edge servers, our framework integrates zk-SNARKs and Merkle trees to ensure data privacy with verifiability. This integration not only maintains privacy but also enables the verification of data origin and integrity. Furthermore, our framework's design for scalability, utilizing edge aggregating servers, addresses the common issue of computational overhead in resource-constrained IoT devices, a significant limitation in many existing techniques. By recording key information such as the Merkle root and proof hash on the blockchain, our approach also ensures immutable and transparent record keeping, enhancing trust and authenticity far beyond what is typically achieved in the current literature. Additionally, our approach to authentication is more performant than similar approaches utilizing ECDSA batch verification.

In essence, the solution shifts the bottleneck to an off-chain system, moving the latency associated with IoT data authentication (signature verification) to this off-chain system. Rollups allow for almost immediate soft transaction confirmation, with the proving time of edge aggregators becoming the main contributor to latency.

Using recursion in our chosen proving system is crucial, as it improves proving time performance. Additionally, proof compression is vital, and without it, the proofs are too large for practical use.

While current results in proving times and proof sizes show that a combination of recursive proving and compression is necessary, we anticipate further advancements and refinements in proof generation. These improvements will likely reduce proof sizes and times even more, allowing us to store the compressed proofs directly on-chain more easily. This would make using smart contracts to verify the proofs easier, enhancing immediacy and integration within the blockchain ecosystem and transitioning our solution from a Layer 2.5 to a Layer 3 system.

Future work could be considered in implementing aspects described in the Security discussion. Another avenue to explore would be other L2 solutions as they mature. While rollups store all transaction data on L1, other L2 solutions like plasma, validiums, and volitions offer different data storage mechanisms, potentially reducing costs even further. Plasma, for example, keeps most data and computation off-chain, except for critical components like deposits, withdrawals, and Merkle roots [34]. Validiums are similar to ZK-rollups but store data off-chain, relying on Data Availability Committees for data storage [22]. Volitions, pioneered by StarkWare, allow applications to switch between ZK-rollup and validium modes, offering flexibility regarding on-chain and off-chain data storage [35]. These solutions present interesting avenues for blockchain-only solutions in IoT data authentication, providing significant cost reductions while maintaining security and integrity.

Author Contributions: Methodology, J.B.B.; Software, J.B.B.; Validation, M.P.; Investigation, J.B.B.; Resources, M.P.; Writing—original draft, J.B.B.; Writing—review & editing, M.P.; Supervision, M.P.. All authors have read and agreed to the published version of the manuscript.

Funding: The research was supported by the Slovenian Research and Innovation Agency and the University of Ljubljana in the research program P2-0425, "Decentralized solutions for the digitalization of industry and smart cities and communities". The research was also supported by the "UL for a Sustainable Society: Lifelong learning and micro-credits" project, co-funded by the Republic of Slovenia, the Ministry of Education, Science and Sport and the European Union—NextGenerationEU.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding authors.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Panchal, A.C.; Khadse, V.M.; Mahalle, P.N. Security Issues in IIoT: A Comprehensive Survey of Attacks on IIoT and Its Countermeasures. In Proceedings of the 2018 IEEE Global Conference on Wireless Computing and Networking (GCWCN), Lonavala, India, 23–24 November 2018; pp. 124–130.
2. Passlick, J.; Dreyer, S.; Olivotti, D.; Grützner, L.; Eilers, D.; Breitner, M.H. Predictive Maintenance as an Internet of Things Enabled Business Model: A Taxonomy. *Electron. Mark.* **2021**, *31*, 67–87. [CrossRef]
3. Al-Ali, A.R.; Zualkernan, I.A.; Rashid, M.; Gupta, R.; Alikarar, M. A Smart Home Energy Management System Using IoT and Big Data Analytics Approach. *IEEE Trans. Consum. Electron.* **2017**, *63*, 426–434. [CrossRef]
4. Kumar, S.; Tiwari, P.; Zymbler, M. Internet of Things Is a Revolutionary Approach for Future Technology Enhancement: A Review. *J. Big Data* **2019**, *6*, 111. [CrossRef]
5. Adi, E.; Anwar, A.; Baig, Z.; Zeadally, S. Machine Learning and Data Analytics for the IoT. *Neural Comput. Appl.* **2020**, *32*, 16205–16233. [CrossRef]
6. Hafid, A.; Hafid, A.S.; Samih, M. Scaling Blockchains: A Comprehensive Survey. *IEEE Access* **2020**, *8*, 125244–125262. [CrossRef]
7. Shen, M.; Liu, H.; Zhu, L.; Xu, K.; Yu, H.; Du, X.; Guizani, M. Blockchain-Assisted Secure Device Authentication for Cross-Domain Industrial IoT. *IEEE J. Sel. Areas Commun.* **2020**, *38*, 942–954. [CrossRef]
8. Liu, B.; Yu, X.L.; Chen, S.; Xu, X.; Zhu, L. Blockchain Based Data Integrity Service Framework for IoT Data. In Proceedings of the 2017 IEEE International Conference on Web Services (ICWS), Honolulu, HI, USA, 25–30 June 2017; pp. 468–475.
9. Zhou, J.; Cao, Z.; Dong, X.; Vasilakos, A.V. Security and Privacy for Cloud-Based IoT: Challenges. *IEEE Commun. Mag.* **2017**, *55*, 26–33. [CrossRef]
10. Barki, A.; Bouabdallah, A.; Gharout, S.; Traoré, J. M2M Security: Challenges and Solutions. *IEEE Commun. Surv. Tutor.* **2016**, *18*, 1241–1254. [CrossRef]
11. Ammar, M.; Russello, G.; Crispo, B. Internet of Things: A Survey on the Security of IoT Frameworks. *J. Inf. Secur. Appl.* **2018**, *38*, 8–27. [CrossRef]
12. Guo, S.; Hu, X.; Guo, S.; Qiu, X.; Qi, F. Blockchain Meets Edge Computing: A Distributed and Trusted Authentication System. *IEEE Trans Ind. Inf.* **2020**, *16*, 1972–1983. [CrossRef]
13. Xu, L.; Chen, L.; Gao, Z.; Fan, X.; Suh, T.; Shi, W. DIoTA: Decentralized-Ledger-Based Framework for Data Authenticity Protection in IoT Systems. *IEEE Netw.* **2020**, *34*, 38–46. [CrossRef]
14. Thantharate, P.; Thantharate, A. ZeroTrustBlock: Enhancing Security, Privacy, and Interoperability of Sensitive Data through ZeroTrust Permissioned Blockchain. *Big Data Cogn. Comput.* **2023**, *7*, 165. [CrossRef]
15. Lee, C.H.; Kim, K.-H. Implementation of IoT System Using Block Chain with Authentication and Data Protection. In Proceedings of the 2018 International Conference on Information Networking (ICOIN), Chiang Mai, Thailand, 10–12 January 2018; pp. 936–940.
16. Xu, R.; Zhou, Y.; Yang, Q.; Yang, K.; Han, Y.; Yang, B.; Xia, Z. An Efficient and Secure Certificateless Aggregate Signature Scheme. *J. Syst. Archit.* **2024**, *147*, 103030. [CrossRef]
17. Fathima, N.; Banu, R.; Ahammed, G.F.A. Integrated Signing Procedure Based Data Transfer Security and Authentication Framework for Internet of Things Applications. *Wirel. Pers. Commun.* **2023**, *130*, 401–420. [CrossRef]
18. Shang, S.; Li, X.; Gu, K.; Li, L.; Zhang, X.; Pandi, V. A Robust Privacy-Preserving Data Aggregation Scheme for Edge-Supported IIoT. *IEEE Trans. Ind. Inf.* **2023**, 1–12. [CrossRef]
19. Kittur, A.S.; Pais, A.R. A New Batch Verification Scheme for ECDSA*signatures. *Sādhanā* **2019**, *44*, 157. [CrossRef]
20. Scaling. Available online: https://ethereum.org/en/developers/docs/scaling/ (accessed on 22 November 2023).
21. Polge, J.; Robert, J.; Le Traon, Y. Permissioned Blockchain Frameworks in the Industry: A Comparison. *ICT Express* **2021**, *7*, 229–233. [CrossRef]
22. Thibault, L.T.; Sarry, T.; Hafid, A.S. Blockchain Scaling Using Rollups: A Comprehensive Survey. *IEEE Access* **2022**, *10*, 93039–93054. [CrossRef]
23. Burgos, J.B.; Pustišek, M. Tackling Trust and Scalability of the Blockchain-Based Shared Manufacturing Concept. In Proceedings of the 2023 17th International Conference on Telecommunications (ConTEL), Graz, Austria, 11–13 July 2023; pp. 1–7.
24. Optimistic Rollups. Available online: https://ethereum.org/en/developers/docs/scaling/optimistic-rollups/ (accessed on 29 December 2023).
25. Zero-Knowledge Rollups. Available online: https://ethereum.org/en/developers/docs/scaling/zk-rollups/ (accessed on 29 December 2023).
26. Thaler, J. *Proofs, Arguments, and Zero-Knowledge*; Now Foundation and Trends: Boston, MA, USA, 2023.

27. Goldreich, O.; Micali, S.; Wigderson, A. Proofs That Yield Nothing but Their Validity and a Methodology of Cryptographic Protocol Design. In *Providing Sound Foundations for Cryptography: On the Work of Shafi Goldwasser and Silvio Micali*; Association for Computing Machinery: New York, NY, USA, 2019; pp. 285–306. ISBN 9781450372664.
28. Petkus, M. Why and How Zk-Snark Works. *arXiv* **2019**, arXiv:1906.07221.
29. Kothapalli, A.; Setty, S.; Tzialla, I. Nova: Recursive Zero-Knowledge Arguments from Folding Schemes. In *Lecture Notes in Computer Science, Proceedings of the Annual International Cryptology Conference, Santa Barbara, CA, USA, 15–18 August 2022*; Springer: Cham, Switzerland, 2022; pp. 359–388.
30. Nguyen, W.; Boneh, D.; Setty, S. Revisiting the Nova Proof System on a Cycle of Curves. *Cryptol. Eprint Arch.* **2023**, *2023*, 969.
31. Boneh, D.; Gentry, C.; Lynn, B.; Shacham, H. A Survey of Two Signature Aggregation Techniques. 2003. Available online: https://networkdls.com/Articles/crypto6n2.pdf#page=2 (accessed on 2 December 2023).
32. Personae Labs Efficient ECDSA & the Case for Client-Side Proving. Available online: https://personaelabs.org/posts/efficient-ecdsa-1/#precomputing-point-multiples (accessed on 5 December 2023).
33. Mud Introduction. Available online: https://mud.dev/introduction (accessed on 17 January 2024).
34. Buterin, V. Exit Games for EVM Validiums: The Return of Plasma. Available online: https://vitalik.eth.limo/general/2023/11/14/neoplasma.html (accessed on 23 November 2023).
35. Volition on Starknet: Your Data, Your Choice. Available online: https://www.starknet.io/en/posts/developers/volition-on-starknet-your-data-your-choice (accessed on 25 November 2023).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Oxford Review of Economic Policy, 2023, **39**, 478–497
https://doi.org/10.1093/oxrep/grad035
Article

Taxing cryptocurrencies

Katherine Baer,* Ruud De Mooij, Shafik Hebous,*** and Michael Keen******

* International Monetary Fund, USA, e-mail: kbaer@imf.org
** International Monetary Fund, USA, e-mail: rdemooij@imf.org
*** International Monetary Fund, USA; CESifo, Germany, e-mail: shebous@imf.org
**** Tokyo College, University of Tokyo, Japan; CERDI, Université Clermont Auvergne, France; Institute for Fiscal Studies, UK; Centre for Business Taxation, UK, e-mail: michael.keen@mail.u-tokyo.ac

We are grateful for helpful comments from Reuven Avi-Yonah, Jeff Hoopes, Tyler Menzer, Joel Slemrod, Finn Tarp, Christophe Waerzeggers, the referee, participants in the conference for this issue, and the discussant of this paper, Alexander Teytelboym. Keen acknowledges support from the Agence Nationale de la Recherche of the French government, through the programme 'Investissements d'avenir' (ANR-10-LABX-14-01). The views expressed here are those of the authors and not necessarily those of the IMF, its Executive Board, or management.

Abstract

Policy-makers are struggling to accommodate cryptocurrencies within tax systems not designed to handle them; this paper reviews the issues that arise. The greatest challenges are for implementation: crypto's pseudonymity is an inherent obstacle to third-party reporting. Design problems arise from cryptocurrencies' dual nature as investment assets and means of payment: more straightforward is a compelling case for corrective taxation of carbon-intensive mining. Ownership is highly concentrated at the top, but many crypto investors have only moderate incomes. The capital gains tax revenue at stake worldwide may be in the tens of billions of dollars, but the more profound risks may ultimately be for VAT/sales taxes.

Keywords: cryptocurrency, virtual assets, tax evasion, tax compliance, bitcoin

JEL classification: E62, H25, H32

I. Introduction

The rise to a contentious prominence of crypto assets has been frenetic, and the pace of innovation involved remains dizzying. From zero in 2008, the market value of crypto assets peaked at around USD 3 trillion in November 2021 (Figure 1); and from Bitcoin, introduced in 2009, have now sprung several thousand other cryptocurrencies. On some estimates, perhaps 20 per cent of the adult population in the US[1] and 10 per cent of that in the UK[2] hold or have held some crypto assets. Use elsewhere is perhaps even more marked, including in some emerging and developing economies: the number of global users has been put at more than 400 million. These developments need to be kept in some perspective: that USD 3 trillion, for instance, was only around 3 per cent of the global value of equities. But the power of developments in crypto assets to disrupt traditional ways of doing financial business, including the collection of tax—and their potential to do more—has been made clear.

To some, these developments presage a brave new world in which people are liberated from oversight by government and reliance on financial institutions, placing their trust instead in cryptographically-protected distributed ledgers, and transactions costs are ultimately greatly reduced. And, beyond this, crypto is seen as the harbinger of wider innovation in the form of decentralized finance that will extend these benefits throughout the financial system. To others, these developments have made crypto markets a 'Wild West'[3] in which criminal activities are

[1] Figures for the US are from surveys whose disinterestedness and quality are not always clear, and orders of magnitude vary widely: towards the low end, InsiderIntelligence (2022) puts at 13 per cent the proportion of crypto holders at end 2022; at the high end, MotleyFool (2022) reports 56 per cent as holding or having held crypto in May 2022.

[2] HMRC (2022*a*).

[3] Gary Gensler, chair of the US Securities and Exchange Commission.

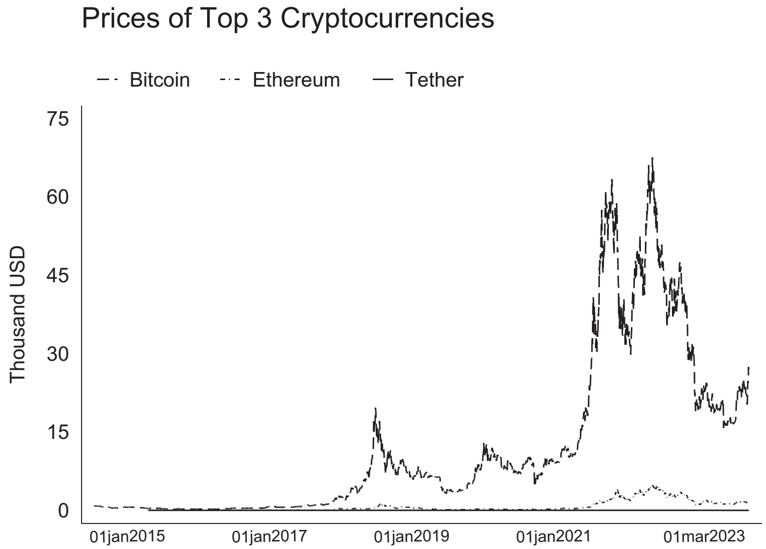

Figure 1: Market capitalization of cryptocurrencies (aggregate and selected types)

Note: The chart shows the total market capitalization of all cryptocurrencies and of a selection of five. Data are from Coinmetrics. In 2014, Bitcoin comprised 95 per cent of market capitalization of all cryptocurrencies. As of November 2022, Bitcoin (40 per cent) and Ethereum (19.5 per cent) are the top two cryptocurrencies in terms of market capitalization, followed by Tether (7.4 per cent). XRP is a cryptocurrency associated with a payment settlement system for financial institutions.

facilitated and poorly informed investors exposed to massive price swings (the USD 3 trillion has now fallen to less than USD 1 trillion), bankruptcies, scams, and frauds (epitomized by the demise of FTX—an exchange platform that also issued its own cryptocurrency—in November 2022). The deepest scam of all, to critics, is that all this is on the basis of assets whose creation creates significant environmental damage and which in many cases have no intrinsic value. In response, advocates might point to the emergence of 'green cryptocurrencies', note that fiat currency also has no intrinsic value, argue that crypto has shown its potential superiority in speed and ease of transactions in the support provided to Ukraine, and assert unknowable benefits from continued innovation.

Regulators face a daunting task in identifying and striking a balance between enabling innovation while securing financial stability and investor protection. For tax authorities, the first-order task is ultimately more mundane, if no easier and no less important: to encompass developments in the use of crypto assets into a well-functioning tax system. Though its importance will differ, that task will remain whatever the future holds for crypto: whether crypto withers or blossoms, the tax system still needs to deal with it.

This paper aims to provide an overview of the issues that the emergence of and likely developments in crypto assets raise for tax design and implementation, with an eye to the implications for the taxation of the rich that is the focus of this issue of the *Review*. The aim is not to provide policy prescriptions, but to set the scene within which decisions must be made and highlight the issues they will need to address.

A large part of the challenge for tax policy and design—beyond that of coming to terms with what remains for many a complex and baffling set of instruments—is that tax systems were not designed for a world in which assets could be traded, and transactions completed, in anything other than national currencies. Incorporating that possibility, however, is more than just a matter of expanding legal definitions (important though in some cases that is). Crucially, the element of anonymity inherent in crypto assets raises issues of enforcement that have long been associated with the use of cash. Those in turn raises issues for the coherence in the taxation of capital income (viewing crypto assets as a form of property) and—less noted, but perhaps ultimately more significant—in the taxation of final sales under the VAT and similar taxes (viewing them as a form of currency). Questions also arise as to whether taxation might have some corrective role to play, complementing regulatory interventions.

Among the most prominent tax concerns, however—and the rationale for a contribution on this topic in the present collection—is the presumption, or suspicion, that crypto assets provide a new and important way for the rich, criminal and other, to evade or avoid taxation. Certainly, crypto has made some people very rich, with, for example, 19 'crypto billionaires' making it to the Forbes List of April 2022.[4] Beyond that, while some of the notably

[4] Forbes (2022). The richest was FTX founder Sam Bankman-Fried, with an estimated net worth at the time of USD 8.7 billion.

rich visibly recoil from crypto,[5] a loosely defined sense that much wealth channelled into crypto escapes proper taxation appears to have become part of the wider mood of dissatisfaction around the taxation of the rich. The concern may have eased since the continuing crypto crash. Auer *et al.* (2022), for instance, suggest that 75 per cent of users have lost money on their Bitcoin investments (which raises its own tax issues around the treatment of losses). How developments in crypto might and should affect the taxation of the rich are among the most important of the challenges they pose—though, as will be seen, to large extent inseparable from still deeper ones.

As well as the challenges there are also real opportunities for tax authorities in the innovations underlying crypto assets. The distributed ledger technology on which they rest, of which blockchain is the most important, is remarkably transparent in the information it contains on the history of transactions, which might ultimately prove valuable for tax administration; and the use of smart contracts (self-executing programmes) within blockchains, for example, might in principle help secure chains of VAT compliance and enforce withholding. The focus here, however, is entirely on the tax challenges associated with crypto assets themselves.

In identifying and taking stock of these, there is relatively little analytical work or empirical evidence to draw on. Within the burgeoning literature on crypto, tax aspects have received relatively little attention. There are compendia on the tax treatment of crypto in various countries,[6] with a useful overview in OECD (2020). And while vast amounts of data are in principle available on transactions in cryptocurrencies, empirical analysis around a complex technology whose central purpose is to leave no tracks is inherently difficult. Experience is accumulating, surveys (doubtless of variable quality) are proliferating, and, importantly, harder evidence emerging from blockchain analytics. The reality, nonetheless, is that this is a technically difficult area in which policy-makers need to act on severely limited information.

To set the scene, the paper first reviews key elements of crypto technologies (with some details and terminology of the mechanics in the Appendix) the ways in which they are traded and used, and by whom. Section III takes up crypto-related issues of tax design. Section IV focuses on the scope for tax evasion that cryptocurrencies offer, and section V turns to the critical issue of tax enforcement. Section VI concludes.

II. Context

This section provides background on the nature and use of the cryptocurrencies that are the main concern of the paper, and on their importance in the taxation of the rich.

(i) The nature of crypto assets

By 'crypto asset' is meant here[7] a 'digital representation of value that relies on a cryptographically secured distributed ledger... to validate and secure transactions'. This produces, without any need for a central authority, a presumptively tamper-free record of transactions in that asset. Categorizing assets within this very wide class in terms of their function—key for their characterizations for tax purposes—is made difficult by both continued innovation and the multiple services that particular assets can provide. That said, one critical tax-relevant dimension along which they vary is between their use for investment purposes and as a means of payment.

- At one extreme are 'security tokens', which are essentially digital representations of conventional financial or other assets. 'Non-fungible tokens' (NFTs), for example, are cryptographically protected representations of unique assets, such as works of art.[8]
- At the other are the central bank digital currencies (CBDCs), which are essentially fiat currency in digital form. Many national governments remain highly cautious on their adoption,[9] but—noting in particular the experiment now under way in China—the general expectation appears to be that, in time, the issuance of CBDCs will become widespread.

Conceptually, the appropriate tax treatment of each of these is straightforward. NFTs are naturally treated as investment assets (though issues of implementation much like those discussed below arise). CBDCs would simply be another form of fiat currency, and so naturally treated as that now is. Depending on their design, CBDCs might have other implications for the tax system: they might, for example, enable tracking of transactions in ways useful

[5] Warren Buffett, most famously, said he would not pay USD 25 for all the Bitcoin in the world (CNBC, 2018).
[6] Such as PwC (2021).
[7] Precise definitions and classifications in the crypto area vary. That which follows is from OECD (2022).
[8] Sothebys, one of the largest multinational art brokers, for instance, is reported to have sold USD 100 million of NFTs in 2021 (Bloomberg, 2021).
[9] A few, however, have already issued CBDCs (in the Bahamas and Nigeria, for example), though it seems with only modest impact.

for tax administration, or even the levying of some form of withholding.[10] That, however, is not the concern here. In what follows we set aside both these categories of crypto asset.

The focus instead is on crypto assets that potentially serve both an investment and a settlement function, and are privately issued. We refer to these as 'cryptocurrencies'.[11] They are also (for now at least) by far the most prevalent form of crypto asset.[12] The essential mechanics of their operation are summarized in the Appendix. Prominent examples include:

- Stablecoins, which are crypto assets that aim to maintain a stable value relative to some specified asset or pool of assets,[13] generally currency(ies), either through some degree of backing in the underlying asset or by algorithmic methods regulating supply. The most prominent are Tether and USDC, both linked to the US dollar. The primary purpose of stablecoins is to serve as a means of payment, and they would be naturally treated as such for tax purposes if they were to achieve the intended stability with probability one. In practice, however, holders are exposed to significant valuation risk, epitomized by the collapse of Terra in May 2022.[14]
- Bitcoin, Ethereum, and similar assets that, while their supply may ultimately be limited (as with Bitcoin), have no intrinsic value. These are sometimes referred to as 'unbacked tokens', or simply as 'non-stablecoins'. While Bitcoin, introduced in 2009, remains the best known and still has the largest share of the crypto market (Figure 1), there are now several thousand alternatives, taking somewhat different forms: Bitcoin operates by proof of work, for example, while Ethereum now works by proof of stake and can incorporate smart contracts; 'privacy coins' offer enhanced anonymity of various kinds; and some cryptocurrencies provide assurances on the greenness of the underlying mining process. Bitcoin is legal tender in (only) El Salvador and the Central African Republic.

It is this intermediate class of assets, and especially non-stable coins, that are the focus here.[15] They raise the most challenging issues of tax design and implementation, combining as they do elements of currency and investment asset.[16] They have been marked too by extensive/notorious price volatility (Figure 2), which creates its own complications in considering their proper tax treatment. The figure also shows that price variability is by no means zero for stablecoins.

A central feature of cryptocurrencies is that, like cash, their use or ownership does not intrinsically reveal the personal or business identity of those involved in a transaction. Holders exercise control through a private 'key' (or address), held in a 'wallet', but transactions reveal (at most) only a public address from which it is encrypted, and

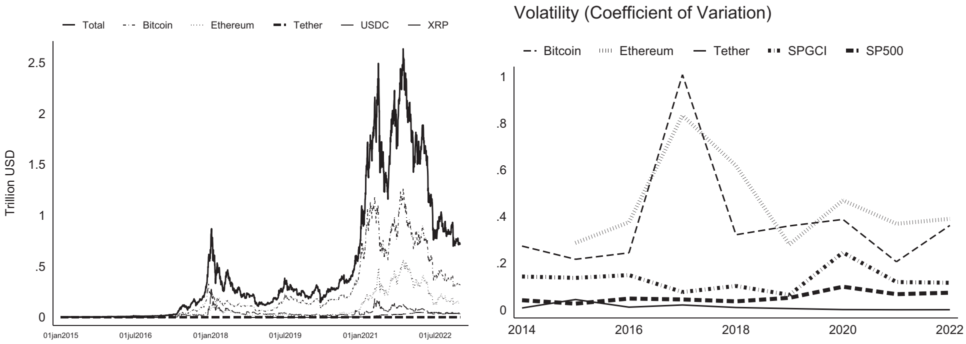

Figure 2: Prices and volatility of cryptocurrencies

Note: The left panel shows the price of 1 unit of a cryptocurrency per USD; Tether aims to maintain its value at 1USD. The right panel shows the coefficient of variation of the three series of prices and, for reference, of the S&P GSCI (a composite index of commodity sector returns) and the S&P500 (stock market index). Underlying data are from Coinmetrics, S&P, and FRED.

[10] See Sarfo (2022).

[11] Also sometimes referred to as 'virtual currencies'. In the US, the Internal Revenue Service (IRS) defines cryptocurrency as 'a digital representation of value that functions as a medium of exchange, a unit of account, and/or a store of value'.

[12] Security tokens have a total market capitalization of around USD 19 billion, far less than the trillions in which cryptocurrencies are measured.

[13] This definition follows IMF (2023).

[14] Only 6 per cent of the liabilities of Tether, for example, are backed by cash (IMF, 2023).

[15] A detailed account of tax issues around stablecoins is provided by Waerzeggers *et al.* (2023). Other forms of crypto assets not considered here include, for example, Initial Coin Offerings, which are crypto assets issued to provide traditional finance, acquired by investors either for resale or to access services of the issuer.

[16] We do not examine the tax issues—or the considerable fiscal risks, emphasized by IMF (2023)—that arise if cryptocurrency is accepted as legal tender.

from which it cannot be inferred. Private addresses are not, in any case, inherently linked to identifiable beneficial owners; and a single user may have (very) many addresses. In this sense cryptocurrencies are generally 'pseudonymous'. (As something of an outlier, Monero, which appears to be the leading privacy coin, is fully anonymous in that it conceals even the public addresses of those involved in a transaction.) Indeed, the creation of difficulty in identifying beneficial owners is one of the primary motivations for the development of crypto assets. It poses obvious problems for tax enforcement, as well as—perhaps more important, in the broader scheme of things—for countering crime, money laundering, and the financing of terrorism.

A second important feature of cryptocurrencies is that, unlike cash, they are remarkably transparent in the sense that these and other details of all transactions on a particular coin are publicly available.[17] This has enabled the development of sophisticated techniques of crypto analytics to analyse patterns of activity and links between participants.

While it appears effectively impossible to recover private addresses from information provided in crypto transactions, the general view seems to be that with sufficient effort IP addresses can often be traced. Again, however, this conveys only limited clues on beneficial ownership, and can be concealed by use of virtual private networks (VPN). Seizures are nonetheless made. In June 2022, for example, the US Department of Justice seized USD 2.3 million of Bitcoins that it assessed to have been paid in a ransomware attack, presumably—though, unsurprisingly, the FBI chose to leave this unclear—obtaining the private key by traditional investigative methods.[18] What several papers do show can be done with some confidence, as will be seen, is to cluster public addresses by the likely nature of their holders: to identify those likely held by exchanges, for example, or those likely held by miners, those involved in ransomware attacks, or those likely having close connection with the darknet.

A third feature of cryptocurrencies that amplifies the difficulties posed by anonymity is their extra-territoriality: transactions reveal no information on the jurisdictional location of those transacting. The ease with which cryptocurrencies can be transacted across national borders immediately casts the tax issue as, in part, one of international cooperation and coordination, with an evident incentive to locate associated activities where tax (and/or regulatory burdens) are light: FTX, for example, was headquartered in the Bahamas.

(ii) Trading in and using cryptocurrencies

There are broadly three ways in which cryptocurrencies may be traded. One is directly peer-to-peer, without the involvement of any third party. The second is through decentralized exchanges,[19] whose purpose is to facilitate such peer-to-peer trades, with customers retaining custody of their private keys. The third is through centralized exchanges,[20] which generally hold their customers' private keys and make transactions on their behalf, charging a commission or fee for doing so. (In this case, as with FTX, the exchange may act much like a bank in trading the crypto it holds, making it potentially vulnerable to a run.) Transactions in centralized exchanges are mostly 'off the chain' in the sense that they are not recorded within the blockchain, the purpose being to avoid potentially sizeable transactions costs[21] (by, for instance, swapping private keys instead). Something more than half of all on-chain transactions take place through decentralized exchanges.[22]

Cryptocurrencies are not yet widely used to purchase goods and services. Press reports suggest that around 15,000 firms globally (400 of them in California) accept Bitcoin and in some cases other cryptocurrencies, including a few household names (such as Microsoft, Overstock.com, Sothebys, and Whole Foods).[23] Information on the extent of actual usage is limited, though it clearly does happen: the crypto analytics firm Chainalysis's (undated) reports put such spending at around USD 15 million per day in early 2021, tiny relative to a daily total of retail sales in the US of around USD 550 billion;[24] one survey suggests that in 2021 about 2 per cent of Americans have used cryptocurrencies to make purchases or money transfers[25] and another, for HMRC (2022*a*) in the UK, reports about 4 per cent of crypto owners as having received crypto for the provision of goods and services.[26] In El Salvador,

[17] See for instance https://www.blockchain.com/explorer.

[18] Department of Justice (2022*b*) simply states that: 'by reviewing the Bitcoin public ledger, law enforcement was able to track multiple transfers of Bitcoin [representing]... the proceeds of the victim's ransom payment, [that] had been transferred to a specific address, for which the FBI has the private key.' Other seizures, to 2017, are listed in the Annex A of Foley *et al.* (2019).

[19] Such as Uniswap (v3) and dYdX.

[20] Such as Binance and Coinbase.

[21] For example, Bitcoin fees averaged USD 2.72 per transaction over the past 3 years; at a median value of Bitcoin transactions of USD 93.61, this implies a transaction fee of almost 3 per cent. At times, the fee has reached USD 60.

[22] Chainalysis (2002*c*). This aligns with the finding in HMRC (2022*a*) that about half of the tax-relevant Group A (described below) used decentralized exchanges.

[23] From Fundera.

[24] Figure for aggregate retail sales from Ycharts.

[25] See Box 2 in Board of Governors (2022). Strikingly, the use of cryptocurrencies for transactional purposes was most prevalent among low-income households and those without bank accounts or credit cards.

[26] See Table 8.3 in HMRC (2022*a*).

where businesses have been required to accept Bitcoin since September 2021, only around 20 per cent in fact do so, and only around 5 per cent of all sales are in this form (Alvarez *et al.*, 2022). Unfamiliarity, high transactions costs, and volatility all impede the routine use of crypto to make purchases. Taxation, as will be seen, can also discourage the use of cryptocurrencies as means of payment. Were these obstacles to fall over time, however, the situation could clearly change; starting from a base of zero, even the low figures in El Salvador are not unimpressive.

The use of cryptocurrencies is not only—or even mainly—a matter for advanced economies. In an index of national penetration produced by Chainalysis,[27] the top four are Vietnam, the Philippines, Ukraine, and India; and the only high-income countries in the top 20 are the US (5th) and the UK (17th). Of the estimated 420 million users, more than one-third are in India, with the highest population shares in UAE (28 per cent), Vietnam (26 per cent), and the US (13 per cent).[28] The reasons for high take-up in emerging and developing countries are not clear. Both Alnasaa *et al.* (2022) and World Bank (2018) find strong positive correlation between the use of cryptocurrencies and indicators of corruption. Though that is suggestive of potential shadiness, more benign interpretations might see the use of crypto as a way to escape corrupt practices and governments that are untrustworthy in their political and/or economic behaviour, and perhaps also as facilitating remittances often important in lower-income countries.[29] In absolute terms, however, the US has the largest crypto market (reaching 16.5 per cent of global crypto value).

(iii) Cryptocurrencies, the rich, and the not-so rich

There are those who have made (and lost) large fortunes from the development of crypto assets and markets. Forbes (2022) lists 19 ‘crypto billionaires’, as noted above, with four among its list of the world’s wealthiest 400 people in 2021 (though that was down from seven in 2019).[30] How much of their wealth is tied up in crypto assets—and so poses problems distinct from those already familiar in taxing the super-rich—is, however, unclear. Harder evidence on those made rich by investing in cryptocurrencies comes from Hoopes *et al.* (2022), who identify all crypto sales in the universe of tax returns in the US between 2013 and 2020, the limitation being that this reveals only those who are both compliant and chose not to continually defer realization. They find 1,245 ‘crypto millionaires’[31] with evidence too that while ‘at least some low-income taxpayers appear to [have experienced] potentially life-changing levels of income via cryptocurrency investments’ many of them were already wealthy.[32]

There are also those who, while deriving their wealth from other sources, are invested in crypto assets. One survey—conducted at the crypto peak in November 2021—suggests that about two-thirds of all Americans with net worth of more than USD 1 million hold some cryptocurrency; and of these, a striking two-thirds hold more than half their wealth in this form.[33] Another survey suggests that in early 2021 there may have been around 100,000 Americans holding more than USD 1 million in crypto assets.[34]

Holdings, moreover, appear to be extremely concentrated. The top 116 addresses, for example, own nearly 16 per cent of all Bitcoins. The link between addresses and individuals, however, is not one-to-one: addresses are held by corporations and intermediaries; they may be jointly owned; and an individual may well hold more than one address. Makarov and Schoar (2021) use algorithms that exploit trading patterns of Bitcoin addresses to distinguish between intermediaries and individual investors. They estimate that the largest 0.01 per cent of individual holders (a total of 10,000) controlled around 5 million Bitcoins, or one-quarter of the total outstanding. This is far larger than the comparable share of equity holdings (and will be even greater if some of these individuals control more than one address).

The apparent propensity of the wealthy to hold significant amounts of cryptocurrency, and the high concentration of holdings (which implies that many are sizeable) means that real issues of fairness and perhaps revenue are at stake in securing their appropriate taxation for income, inheritance, and wealth taxation—including, not least, the treatment of losses.

It is also important, however, to remember that it is by no means only the wealthy who hold crypto assets: in the US, survey evidence suggests that 30 per cent of holders have annual income below USD 50,000;[35] Hoopes *et al.*

[27] See Chainalysis (2022*b*). This is a composite reflecting five types of cryptocurrency services and giving higher weight to countries with lower income per capita.

[28] According to TripleA.

[29] That said, Alnasaa *et al.* (2022) find no significant relationship with remittances or domestic inflation rate when controlling for governance.

[30] Forbes (2021).

[31] Meaning those with cumulative reported gains on cryptocurrency of USD 1 million or more over this period.

[32] Their average wage income was over USD 360,000.

[33] MotleyFool (2021).

[34] CBS (2021).

[35] Board of Governors (2022).

(2022) find even greater prominence of those on lower incomes: over half have taxable income of less than USD 40,000. For the UK, HMRC (2022*a*) finds that 85 per cent of crypto holders have income below £50,000.

III. Cryptocurrencies and tax design

This section considers the key questions of policy that arise in framing and assessing the tax treatment of cryptocurrencies, deferring until later the issues of administration with which they intersect. Following the chain of events through transactions in and the creation of cryptocurrency (Figure 3), issues arise in relation to both income taxation and VAT/sales taxes;[36] there may also be scope for purely corrective taxation. Countries' current practices in these areas are diverse, in many cases await clarification, and are generally in a state of flux.[37]

The natural principle to apply in approaching these design issues—externalities aside, for the moment—is that of neutrality: taxing cryptocurrencies in the same way as comparable traditional instruments. For miners, for example, there seems no reason to treat income from fees and the generation of new coins differently from other business income, unless some specific (dis)incentive is intended. Application of neutrality principles to the treatment of cryptocurrencies is made difficult, however, by their dual nature: as investment assets and as a medium of exchange.

(i) Income taxation

Corresponding to these two functions, there are two main ways in which cryptocurrencies might be classified for income tax purposes: as property (like shares, or bonds) or as (foreign) currency. The implications of the difference will depend on domestic rules, but can be highly material. For instance, many countries exempt individuals' capital gains on foreign currencies (Cnossen and Jacobs, 2022) whereas classification as property usually gives rise to capital gains tax (although important detail on, for instance, any ring-fencing of losses, exempt amounts, and variation of rates with holding periods, will differ). In the US, for example, the characterization of cryptocurrencies as property means that capital gains are in principle reportable on all transactions, with lower than the ordinary income tax rate applying if held for more than 1 year; had they instead been characterized as currency, gains would be taxable as ordinary income but only on gains over USD 200. Similar difficulties arise elsewhere, with the treatment of cryptocurrencies as property requiring calculation of gain or loss on every transaction. The obligations this imposes on small users are potentially extremely burdensome, and a significant obstacle to the routine use of cryptocurrencies to acquire goods and services.[38]

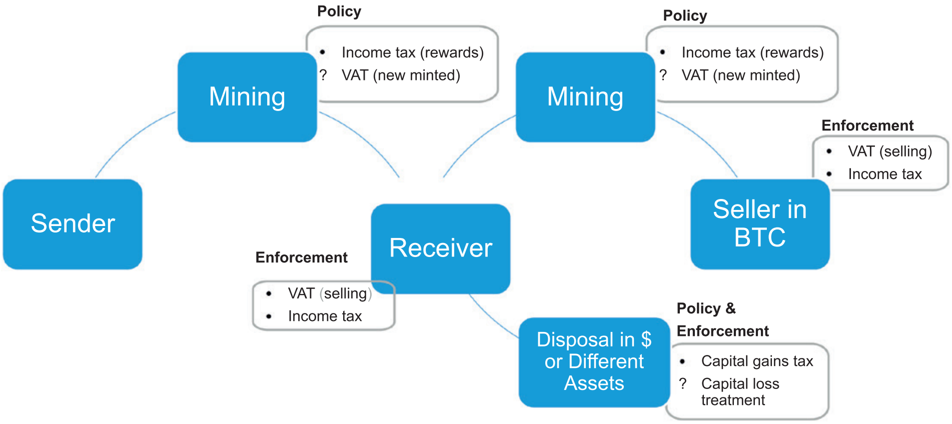

Figure 3: An illustrative chain of events.

Note: This diagram illustrates taxable events from the circulation of a cryptocurrency—taken here to be Bitcoin (BTC)—highlighting their particular tax policy and administrative challenges. The sender, via miners, purchases a service from the receiver using BTC, and the receiver has the options to either dispose the BTC or purchase a service with the BTC. A '?' indicates a particular need for policy/legal clarity. Not explicitly depicted here is that these transactions can be peer-to-peer (P2P), or via decentralized or centralized exchanges; this does not affect the policy treatment but affects tax enforcement capability (P2P being most difficult, followed by decentralized, and finally centralized exchange).

[36] There are, of course, implications for other taxes, including on wealth, gifts, and inheritances, that for brevity we do not pursue here.

[37] For a systematic account, see OECD (2020).

[38] See, for instance, Wiseman (2016), who suggests that users be entitled to self-select as investment or transactional users (or at least that a *de minimis* personal exemption be introduced). Waerzeggers *et al*. (2023) highlight this difficulty for stablecoins in particular, given what is clearly their primary purpose of serving as a means of payment, and explore further the complications associated with the differences between the many types of stablecoin.

There is perhaps a third possibility. Some have drawn analogies between the holding of cryptocurrencies and gambling, with the apparent implication that it should be taxed in the same way: see for instance Panetta (2023). This would have implications not only for income taxation but for VAT and sales tax (with acquisition being treated as a stake), which treat gambling in complex and diverse ways.[39] The aptness of the analogy, however, is unclear: about half of respondents in HMRC (2022*a*) report holding cryptocurrency 'just for fun', but Hoopes *et al*. (2022) find that crypto sellers look much like everyone else in terms of their reported gambling income.

In practice, the most common approach appears to be to treat cryptocurrencies for income tax purposes as property and subject to the corresponding capital gains tax rules. That still leaves room for a wide variety of treatments. Several countries, including in Europe, Malaysia, and Singapore, either do not tax capital gains from financial assets or exempt gains after a rather short holding period.[40] Portugal, which has tried to position itself as crypto-friendly, has specifically exempted gains on crypto holdings, though this now applies only to holdings of over 1 year; El Salvador still has an outright exemption.[41]

One special case of note is that of India. There crypto assets are in a regulatory limbo: neither illegal nor, strictly speaking, legal. Nevertheless, the government has implemented a bespoke tax regime specifically aimed at taxing at 30 per cent gains and/or income from trading in 'virtual digital assets' (VDAs), meaning cryptocurrencies, NFTs and similar tokens, and other assets that it may specify. This is accompanied by a 1 per cent surcharge on the transfer of any VDA.[42]

Unsurprisingly, the tax treatment of capital losses from holding crypto is coming to receive considerable attention. The question is whether the use of these should be subject to limitations beyond those generally imposed: in the US, for example, capital losses can be offset only against capital gains. By imposing tough limits, government can, if it wishes, reduce both the attractions of trading in crypto and the extent to which it shares in the risks involved. Nguyen and Maine (2023), for instance, argue for allowing crypto losses to be offset only against crypto gains. Inventing special rules for crypto in the aftermath of massive losses does, however, risk proving a hasty and unprincipled response to unusual circumstances. That said, one can argue that any limits on the use of capital losses are ultimately a matter of pragmatism rather than inherent in the core principles of income taxation. What then becomes important are the considerations behind the pragmatism. Discouraging use, for example, might also or alternatively call for a differentially high rate on crypto gains; and preserving revenue achieved, perhaps more fairly, by directly limiting the amount that can be offset rather than what the offset may be against.

(ii) VAT and sales taxation

The use of cryptocurrencies should pose no great difficulty of principle for the core structure of these taxes, since—with barter transactions in mind—these are commonly couched in terms of supply being made not for legal tender but for 'consideration', a term broad enough to encompass crypto assets. (Practical difficulties in applying this may well arise, however, some of them touched on later, from price volatility (which can place particular pressure on verifying precisely when transactions occur), scope for fraud, and incorporation into cross-border rules.) To ensure that the acquisition of cryptocurrency for fiat money is not in itself subject to VAT, several countries (including Australia, Japan, and South Africa) provide an explicit VAT exemption;[43] in the EU, the Court of Justice held in 2015 that VAT should not be applied to such transactions.

The VAT treatment of the fees and newly minted cryptocurrencies received by miners also requires a clear policy stance. In principle, there seems no reason why—again except by way of creating a deliberate (dis)incentive[44]—these should not be fully liable to VAT, with a corresponding right to credit of VAT charged on inputs. While that is generally recognized as good practice, many VATs in practice exempt fees for financial services. This will result in over-taxation of business use of cryptocurrency (because of miners' unrecovered input VAT) and under-taxation of individual use.

(iii) Externalities

Several types of externality might be associated with the use of cryptocurrencies, and indeed these are reflected in the calls in many countries for their more effective regulation and, in some (including China, Egypt, Bolivia, and Bangladesh), the outright prohibition of trading in or mining cryptocurrencies.[45] Beyond those externalities

[39] See, for example, Clotfelter (2005).

[40] Auer and Tercero-Lucas (2022) find that owners of cryptocurrencies increasingly tend to hold for longer periods.

[41] Where cryptocurrencies are brought into capital gains tax, the question—partly conceptual, but above all practical— also arises as to whether the exchange of crypto for other virtual assets should be deemed a realization of gains/losses or a step-up/down in the asset's base for taxation later upon realization: see Avi-Yonah and Mohanad (2022).

[42] Above an annual threshold of around USD 600.

[43] OECD (2020).

[44] The question of whether to allow crediting of electricity costs is taken up below.

[45] Mining seems, nevertheless, to have lingered in China at least into 2022.

conventionally addressed through regulatory measures aimed at ensuring financial stability, protecting consumers, and countering criminality, however, some may be associated directly with the use of cryptocurrencies themselves.

The analogy with gambling mentioned above, for example, points to possible problems of self-control of a kind that can rationalize corrective taxation. Extensive substitution of cryptocurrencies for national currencies ('cryptoization') might undermine the tools of macroeconomic management, significantly reducing the effectiveness of monetary policy or capital flow measures—with possible implications for the functioning of the international monetary system. For both problems, the possibility arises of correction through some form of tax on transactions in cryptocurrencies, akin to the financial transactions taxes that have been adopted and even more often proposed for traditional instruments (including to reduce excessive price volatility,[46] of a kind that many also associate with cryptocurrencies). It might also be that, pending more effective regulation, using the tax system to discourage trading could in principle serve as a (very) second-best stop-gap device to counter risks to financial stability and dampen risks to poorly informed investors. The 1 per cent transfer tax in India might indeed be seen as a pioneering step addressed to these purposes. But whatever the conceptual merits of a crypto transactions tax might or might not be—and there are counterarguments in the unknown benefits of fostering innovation in crypto—implementation is problematic, for reasons similar to those highlighted in section V: while national application to the subset of transactions through centralized domestic exchanges (and/or miners) may be feasible, this might simply drive transactions into decentralized or peer-to-peer forms, or offshore. Similar arguments might nonetheless warrant less dramatic measures within existing structures, such as denying or limiting loss offsetting under the capital gains tax.

The most compelling case for feasible corrective taxation, however, is environmental. Proof-of-work consensus mechanisms (such as that behind Bitcoin) require considerable energy, as they rest on finding a solution to a complex mathematical problem by making an enormous number of guesses. The associated carbon emissions are cause for considerable concern: Hebous and Vernon (forthcoming), for instance, estimate that in 2021 Bitcoin and Ethereum used more electricity than did, for instance, Bangladesh or Belgium, generating 0.28 per cent of global greenhouse gas emissions.[47]

Awareness of this problem is now quite widespread, and reflected in the explicit marketing of some cryptocurrencies as 'green'. Voluntariness alone, however, is unlikely to provide a complete solution.[48] By the usual arguments, externalities from mining-related carbon emissions are best addressed within a general carbon tax, which would automatically internalize the costs of the energy-heavy proof-of-work verification mechanisms. In the absence of a carbon tax, however, there is a case for more targeted tax measures. In March, the Biden administration proposed a 30 per cent tax on miners' electricity use, though (for now at least) with no differentiation to reflect the carbon-intensity with which it is generated. Kazakhstan (an important location for mining) introduced a similar tax at the start of 2023, with a reduced rate for those using renewable sources.[49] Absent any additional tax of this kind, a less efficient but nonetheless meaningful measure might be to limit or deny income tax deductions for energy costs incurred in their mining activities, and/or similarly that (if not VAT exempt) there be no credit for input VAT on energy costs.

IV. Evasion and revenue potential

From their outset, a paramount concern with cryptocurrencies and crypto assets more generally has been the appeal of their anonymity properties in facilitating criminal activities. And there is no doubt that their criminal use is extensive, both in the large seizures that there have been (the largest, in February 2022, being of Bitcoin valued at USD 3.6 billion)[50] and in the sizeable (though short-lived) price responses to them. Mention has already been made of the strong negative cross-country correlation between usage of Bitcoin and various indicators of institutional quality/control of corruption. Specific areas of concern include both 'traditional' crimes—money laundering, trade in drugs and other illegal goods and services, financing of terrorism—and newer ones that draw on similar digital

[46] Whether they in fact do so is another matter, and likely to depend on market characteristics, such as thickness and the prevalence of institutional investors. For an overview of the issues around financial transactions taxes, see Matheson (2012).

[47] One might be tempted to argue further that, irrespective of any environment impact, the resources that these procedures use are in any case in effect a deadweight cost: the mathematical problem is of no substantive importance, and is required only to impose a cost that discourages the sending of dishonest messages. Abadi and Brunnermeier (2022) show, however, that some such costs are required if a consensus mechanism is to have desirable properties. Short of arguing that crypto in itself is inherently worthless, this resource use thus has some value.

[48] In September 2022, Ethereum moved to a proof-of-stake mechanism that is far less damaging to the environment, but there are no signs that Bitcoin and others will abandon the proof-of-work consensus any time soon.

[49] By way of context, with shares of 35 and 18 per cent respectively, the US and Kazakhstan account for more than half of all mining; Russia follows at 11 per cent.

[50] This related to theft from Bitfinex, a cryptocurrency exchange: see Department of Justice (2022*b*).

skills, including online frauds and ransomware attacks. Tax evasion is commonly included in this long list, but, perhaps unsurprisingly, often rather low down. And, indeed, it is easier to get at least some direct handle on the use of crypto for criminal activity in general than for tax evasion in particular. We consider each in turn.

(i) Crime and crypto

More is known about the use of crypto for what may be more serious crimes than its use for tax evasion. This is because the nature of the technology—with the entire history of transactions on a blockchain and public keys all public information—turns out to provide meaningful clues as to the extent of hard-core criminal activities.

With data on the universe of (over 600 million) transactions in blockchain from its inception in 2009 until April 2017, Foley *et al.* (2019) start by identifying addresses associated with seizures and trading on the darknet.[51] Building on this, they go further and estimate a wider population of those likely to be engaged in illegal activities, both by building clusters based on trades with those identified as 'directly' illegal and estimation based on their characteristics (such as the use of 'tumbling'[52] and other measures of obfuscation). Their final estimate is that in 2017 around 25 per cent of all Bitcoin users were engaged in criminal activity, accounting for around 23 (respectively, about 17)[53] per cent of all transactions by number (by value), and holding around half of all Bitcoins. In dollar terms, this means transactions of USD 76 billion and Bitcoin holdings of USD 7 billion in 2017. These are large enough figures to conclude that 'a significant component of [Bitcoin's] value as a payment system derives from its use in facilitating illegal trade'.[54] Against this, however, Chainalysis, a specialist crypto analytics firm, arrives at much smaller numbers for the relative scale of illegal activities enabled by cryptocurrencies. For that same year of 2017, Chainalysis (2022*a*) estimates that 'illicit addresses' accounted for around 1.4 per cent of all transactions, and received payments of around USD 4 billion. Makarov and Schoar (2021) arrive at a similarly modest figure, putting illegality and gambling at under 3 per cent of the total.

These are very substantial differences. Makarov and Schoar (2021) attribute them largely to differences in the denominator, with their own estimate, unlike that of Foley *et al.* (2019), including exchange related and similar transactions: these account for about 90 per cent of total volume. While this helps reconcile the figures, since Makarov and Schoar (2021) regard these transactions as not 'economically meaningful', the implication would nonetheless be that illegal transactions are a sizeable share of Bitcoin transactions that relate to real activity.

There is, though, consensus that while crypto-enabled crime is growing rapidly in absolute volume and value, it is declining in relative terms.[55] It also seems to be agreed, albeit more asserted than demonstrated, that criminality continues to rely more heavily on traditional financing means, including cash.[56]

What is also clear is that cryptocurrencies continue to innovate towards increasing anonymity. Monero, for example, operates on the blockchain but has the feature that public keys, transaction details, and other information are not publicly revealed. These are increasingly seen as the cryptocurrency of choice for serious crime. In response, in 2020, the IRS issued a request for proposals to enhance capacity to trace transactions on Moreno and other privacy keys.

(ii) Evasion and crypto

Since the proceeds of illegal activities are generally taxable, estimates of illegality of the kind above will encompass some degree of tax evasion. For the serious crimes that have been the primary focus of concern and investigation, however, tax evasion is likely a by-product rather than a primary motivation; indeed, the purpose of the extensive money laundering in cryptocurrency is to make illegal gains appear legal, and even perhaps consequently subject to some tax.

From the narrower tax perspective, three issues are critical. The first is the qualitative nature of the incentives to use crypto as a means of evading taxes on what are otherwise legal transactions: the provision of legal goods or services, for example, or payment of salaries. The second is the extent to which cryptocurrencies are, or might be, used to such an end. The third is the extent to which taxes related to the generation of or trading in crypto assets themselves—earnings from mining, for example, or gains on their sale—are properly paid.

On the first question, crypto technologies do not in themselves fundamentally alter the considerations shaping decisions on engaging in tax evasion. Their essence remains as set out in the literature that begins with Allingham

[51] For 2017, about 6 per cent of all Bitcoin users are placed in this category, accounting for around one-third of all transactions.
[52] Tumbling involves mixing funds from different sources and sending on to distinct addresses; it is intended to make tracking harder.
[53] From inspection of their Figure 6B.
[54] Foley *et al.* (2019), p. 1802.
[55] This was true over the sample period of Foley *et al.* (2019) (and see also Tasca *et al.* (2022)); for the period since, Chainalysis (2022*a*) estimates that the amounts flowing to 'illicit' addresses have about tripled, to around USD 15 billion, but that as a share of all transactions they have fallen by 90 per cent.
[56] See for instance Europol (2021).

and Sandmo (1972): balancing the tax saved in the event that evasion succeeds against the losses suffered in the event that it does not, with the likelihood of each depending on the probability of detection. In that canonical setting, the distinctive feature of crypto, arising from its anonymity, is naturally thought of as a particularly low probability of detection and hence particular appeal as a device for evasion. But there may be further considerations entering the evasion calculus.

One is that transactions costs will matter, both as to whether to evade at all and in selecting the means of doing so. In this respect, where the balance of advantage lies between crypto and, the obvious comparator, cash, is not a priori clear-cut (and may change over time). Crypto may, for example, offer some savings in transactions costs relative to cash,[57] for which existing channels can sometimes be more expensive; the estimated fee for a $200 remittance, for instance, is 5.7 per cent compared to 1.4 per cent for a Bitcoin transaction (Beck *et al*., 2022). But its use and exploitation to enhance anonymity may require skills that are costly to acquire. Over time, the latter obstacle to the use of crypto (leaving any regulatory restrictions aside) will surely decrease.

Another, more distinctive consideration, is the risk implied by the high price volatility of cryptocurrencies.[58] Differentially higher risk of fraud or theft in using cryptocurrencies would have similar effect. Akin to exchange rate risk, the primary consequence is likely to be to discourage the use of crypto for evasion purposes. Limiting it may lead risk-averse evaders to cash out more quickly than those holding for investment purposes, a trait that the studies cited above suggest is associated with criminal use of crypto.

There is almost no hard evidence, even anecdotal, on the second, quantitative question as to the extent of crypto-enabled tax evasion. In India, the authorities seized nearly USD one billion in evaded Goods and Services Tax (GST) from local exchanges in May 2022.[59] And, in the UK, HMRC seized crypto assets and NFTs apparently intended to set up a VAT fraud.[60] It is interesting, and perhaps telling, that both episodes relate not to income taxation but to the VAT. Being left, however, with even less firm evidence on the possible extent of crypto-related evasion than there is for more traditional forms, we explore the possibilities further in the next subsection.

On the third question, there is some information on the converse of evasion: tax paid. For the US, the results of Hoopes *et al*. (2022) imply that around 1 per cent of all returns in 2020 reported some sales of crypto. That is well below the 10–20 per cent or so of American adults suggested by survey evidence to have held crypto around then; but of course some compliers may simply have chosen not to realize. There are also signs of at least some degree of compliance in the UK: a survey conducted for HMRC (2022*a*) reports that 45 per cent of crypto owners thought they might be subject to capital gains tax, and 34 per cent believed that they had a good understanding of the rules. It also reports that nearly 30 per cent of owners had sought their guidance on the tax treatment of crypto—most, presumably, intending to comply rather than seeking clues on how to evade. While that might suggest fairly widespread understanding, many holdings were so small that any gain would almost certainly be below the exempt amount.[61] Cluster analysis identifies only one group (‘A’) as having tax-relevant gains (or losses)—within which,[62] 90 per cent believe they have a good understanding of capital gains tax and 72 per cent have seen HMRC guidance. Knowing one’s tax obligations is not the same, of course, as intending to comply with them; and it is noticeable that those in Group A are also more likely to trade outside centralized exchanges, including peer-to-peer, in ways that are harder to monitor. More generally, it seems unlikely that those determined not to comply will respond to a survey on this topic. While there is thus some comfort in these results, it is far from complete.

Cong *et al*. (2022) look for signs of tax compliance by exploring the extent to which US-based crypto owners harvest tax losses around year end; that is, sell crypto and immediately repurchase so as to realize losses while leaving their holding unchanged—a transaction that would be hard to rationalize other than as one of tax planning. (For traditional securities, such harvesting is restricted by disregarding repurchases within 60 days of the sale; as one of the grey areas in the precise details of the taxation of crypto, however, this restriction was not believed to apply to crypto.)[63] Using both proprietary data from 500 large retail traders and data from 34 exchanges,[64] Cong *et al*. (2022) find that such crypto transactions do occur and, moreover, increased following public statements by

[57] IMF (2023) is sceptical on this, at least at present, except perhaps for some small cross-border payments of the kind in the example that follows.

[58] To a lesser degree, of course, for stablecoins.

[59] ITR (2022). Traders in cryptocurrencies are subject in India to 18 per cent GST.

[60] See Financial Times, https://www.ft.com/content/3695637a-9f8f-4001-9e3b-c65adffef4db, and Cointelegraph, https://cointelegraph.com/news/uk-tax-authority-makes-first-nft-seizure-in-vat-fraud-case.

[61] At the time, £12,500. In principle, they might also have gains on other assets.

[62] Figures that follow are from HMRC (2022*a*, Table 8). Of those in Group A (marked by higher incomes and much more frequent trading), 64 per cent were estimated to have a taxable gain on crypto alone; in the other three groups, that figure is no more than 1 per cent.

[63] The administration has now made clear its intention to ensure that it will.

[64] The latter may be more compelling since, as the authors note, their willingness to provide data may suggest that traders intend to be broadly compliant.

the IRS highlighting tax obligations on crypto transactions and its intention of targeted enforcement efforts. The conclusions to be drawn from this, however, are somewhat mixed: while there is a degree of compliance, the impact of policy statements by the IRS suggests that, whether wilful or the product of ignorance, there is, or at least has been, a consequential element of non-compliance.

(iii) Revenue potential

All this leaves little sense as to the amount of revenue at stake, whether for collection or evasion. Perhaps coming closest to this goal is the work by Thiemann (2021). Using data on Bitcoin transactions provided by Chainalysis, linked probabilistically to country of user (by information on web traffic flows to platforms and other clues such as time difference) this arrives at estimates of accrued and realized capital gains by EU residents. While information on capital gains tax actually paid on these transactions is unavailable, this enables rough estimation of the tax due in principle—which in turn can be thought of as an upper bound on the amount of tax evaded. And that amount is put at EUR 850–900 million in 2020. It is hard to scale this relative to revenues from the taxation of personal capital gains tax in the EU, which many countries do not report. But, for example, Thiemann (2021) suggests that this is about 0.3 per cent of total property tax revenue in the EU; and it compares to capital gains tax revenue in the UK alone (not included in the sample) of about EUR 12 billion.[65]

Beyond this, so great is the ignorance in this area that even the crudest back-of-envelope calculations may be helpful. In this spirit, one approach is to treat cryptocurrencies as an investment asset, and apply some assumed rates of return and of taxation. Suppose, for instance, a total crypto market capitalization of USD 1 trillion[66] (which, for scaling, might be compared with well-known estimates that global 'hidden wealth' is around USD 7 trillion).[67] Assuming a rate of return of 5 per cent—having in mind returns closer to normal than in the past—and a tax rate of 20 per cent (brushing aside the complexities and diversity of national tax regimes), the implied total tax due is USD 10 billion. Taking instead the peak market valuation (in November 2021) of USD 2.6 trillion, this becomes an annual USD 26 billion.

These may all seem small numbers in the global context: even that latter, for instance, is only around 1 per cent of worldwide revenue from the corporate income tax (CIT). However, while an assumed return of 5 per cent may approximate some kind of steady state for cryptocurrencies, there is little sign of an approach to such normalcy, and, to the contrary, the past has been characterized by remarkable volatility. Looking back at the past 2 years can give a broad idea of how much revenue has been at stake. Market capitalization of crypto assets in 2021 went from USD 752 to USD 2,368 billion; in 2022, it dropped to USD 836 billion. These numbers are close approximations of the accrued capital gain/loss (the volume of Bitcoins, for instance, rose by only 2 per cent and these newly minted Bitcoins would be subject to income tax too). Again assuming a tax rate of 20 per cent, the tax revenue from an accrual-based capital gains tax would thus have been a hefty USD 323 billion in 2021, or around 12 per cent of global CIT revenue. If only one-third of these gains were realized, revenue would still be around USD 100 billion. In 2022, on the other hand, if losses had been fully offset against other incomes, the reduction in tax revenue loss would have been of a similar magnitude.

Given the high concentration of holdings noted above, the potential tax on the gains (and offsets on the losses) of the largest holdings would be correspondingly huge: in 2021, for example, the implied tax on the accrued gains of those 116 largest addresses would be around USD 17 billion. In the 'normal times' scenario, it would be around USD 1.4 billion.

One might also consider——not as a recommendation, though nonetheless with the corrective considerations discussed above in mind—the revenue that would be raised by applying to cryptocurrencies a financial transactions tax of the kind sometimes proposed or applied to trading in securities. At, for example, the rate of 0.1 per cent on securities trading in the still-lingering proposal of the European Commission (EC, 2011), revenue, if applied to all crypto transactions, which amounted to USD 15.8 trillion in 2021,[68] would be around USD 15.8 billion.[69]

An alternative approach is to focus on the use of cryptocurrency as a means of payment. A difficulty here is that some transactions would in principle be fully taxable (such as purchases of goods and services or property by final consumers)[70] others (purchases of business inputs, including salaries) would be wholly or partly deductible by the

[65] From OECD Revenue Statistics (https://stats.oecd.org/), accessed 22 November 2021; at an exchange rate of £1=€1.2.

[66] https://coinmarketcap.com/charts; accessed 11 November 2022.

[67] See Zucman (2013) and Alstadsæter *et al.* (2019).

[68] Chainalysis (2022*a*). Trading volume in 2022 is around 30 per cent lower than in 2021.

[69] At the still lower rate of 0.01118 per cent applied to certain transactions in Brazil, it would come to only USD 1.8 billion (KPMG, https://kpmg.com/us/en/home/insights/2021/09/tnf-brazil-tax-on-financial-transactions-iof-increased-rates.html).

[70] Use for this purpose may not be trivial: HMRC (2022*a*, p.6), reports 9 per cent of crypto owners as having been paid for work in crypto assets. While, as noted above, only 4 per cent indicate having received crypto for the provision of goods and services, for Group A this rises to a strikingly high 26 per cent (HMRC, 2022*a*, Table 8.3).

purchaser. Even generous allowance for the latter, however, suggests that large amounts could be at stake. Suppose, for example, that all cryptocurrency transactions were in the form of a VAT chain, with final sales accounting for 5 per cent of all transactions by value. With total transactions of USD 15.8 trillion in 2021, at a VAT/sales tax rate of 15 per cent the tax due would be a massive USD 118.5 billion. It may be, of course, that many of the current transactions relate to serious crime, and in that sense are likely an order of magnitude harder to recover than tax on legal activities. But if only 2 per cent of transactions were for legal final sales, the implied revenue would still be a sizeable USD 47.4 billion.

These calculations are scandalously simplistic, with more caveats—the assumption, for example, that tax changes would not affect volumes or values—than are worth listing. And of course, since some of that potential revenue is presumably already being collected, they do not indicate how much tax is being evaded. Some sense of that is provided by the estimate of the US Joint Committee on Taxation (2021) that revenue in the first year of operation of the new crypto reporting requirements described below—presumably additional to what would have been raised though self- or other reporting, and perhaps affected by the cumulation of losses since 2021—would be USD 1.5 billion, rising to USD 4.6 billion in 2031. This is around 1 per cent of total (federal, state, and local) revenue from individual capital gains tax in 2020.

All this does suggest some lessons. One is that the revenue at stake worldwide is plausibly in the tens of billions of dollars; perhaps even, if cryptocurrencies were to perform strongly, in the high tens. How much of that is plausibly recoverable is another matter: even leaving aside the problems of detecting evasion associated with pseudonymity, controlling serious criminality lies far beyond the reach of tax administrations. A second is that a large part of the revenue at issue relates to large and wealthy holders of cryptocurrencies. And a third is that, in revenue terms, the use of crypto as a currency for legal transactions rather than for investment purposes might perhaps become a still greater concern. Much of the attention that has been paid to cryptocurrency in the tax literature has been around income tax issues; it may be, however, that it is in relation to VAT and sales taxes that the most significant issues will arise.

V. The heart of the matter: implementation

A prerequisite for effective and efficient tax enforcement is a clear and complete statement of the rules to be implemented. Even in some of the most advanced countries, however, important detail in the tax treatment of crypto assets remains unclear.[71] Indeed so rapid are developments in the area, and so hard have they proved to accommodate within pre-existing legislation, that the process of tax rules reactively trying to catch up with innovations in technology and associated financial operations[72] seems likely to continue for some time.

(i) The implications of anonymity

The fundamental obstacle to tax enforcement in relation to cryptocurrencies is the element of anonymity. The novelty of the problem was nicely expressed by the discussant of this paper (somewhat reformulated): in the old days, the tax authorities' problem was that it knew who you were, but not your transactions; with crypto the problem is that it knows your transactions but not who you are.[73] The aphorism conveys the important element of truth that the use of cryptocurrencies raises problems not so much in identifying transactions but in linking them with specific entities, and in discerning their purpose, of a kind never before faced. Nor is this an accidental by-product of the development of cryptocurrencies. Their creation has been driven precisely by an intention (whether for libertarian or criminal purposes) to provide ways to undertake financial transactions without involving either government itself or any central authority that would be in a position to provide the government with third-party information. With no intrinsic motivation to self-report crypto transactions (except perhaps when they make a loss), overcoming pseudo-anonymity is the core problem that tax administrations are now trying to address.

It is worth pausing, however, to reflect on the significance of anonymity for tax implementation. In itself, anonymity—meaning an inability to link transactions with specific individuals or legal entities—is not in itself fatal to any form of taxation. A single rate transactions tax requires no information on taxpayers' identities; what impedes its anonymous implementation in the blockchain case is an inability of the tax authorities to insert themselves into the chain. Similarly, implementing a truly flat rate income tax does not require the identification of taxpayers; it does though require information, also not available on the blockchain, on the nature of transactions (to pick out,

[71] As, for instance, with the issue noted above of whether wash rules in the US apply to cryptocurrencies.

[72] A recent example is the consultation on the tax issues related to staking and lending crypto assets launched by HMRC in July 2022 (HMRC, 2022*b*).

[73] The idea, though not quite the same turn of phrase, is in Baronchelli *et al.* (2022).

for example, a payment of interest from the purchase of a commodity). More complex tax structures (nonlinear income taxes, or VATs involving both credits and output tax) require some means to identify distinct transactions with the same individual so as to aggregate over them, including transactions conducted by means other than cryptocurrencies. In the limit, if all transactions were conducted in cryptocurrencies, and individuals or firms each had a unique digital identifier, one could conceive of sophisticated tax systems being implemented, with the use of smart contracts, entirely on the blockchain—and this in principle would not require the tax authorities to identify the real individuals and firms behind those identifiers. Privacy in this respect could be fully respected and, in principle, the dystopian prospects some see as an ultimate result of CBDCs[74] avoided. But whether governments would abstain from building in some ability to identify specific, real individuals can be doubted.

For the foreseeable future, however, the challenge for tax authorities is the less profound but very demanding one of finding ways to accommodate the pseudo-anonymity of cryptocurrencies within systems not designed to do so.

(ii) Dealing with pseudonymity

The good news for tax authorities, and regulators, is that—contrary to the vision of the original crypto designers—a core role has emerged for centralized institutions of various kinds in the transacting of crypto assets, notably exchanges through which they are bought and sold. Such institutions are in a position to obtain information on ownership, and so are at the core of current efforts, perhaps somewhat belatedly and certainly still incompletely, to obtain useful third-party information that can be shared with tax authorities.

The use of intermediaries to either acquire or cash out crypto for fiat currency or other traditional instruments is a natural point for tax authorities to acquire information. An important step to this end is ensuring that anti-money laundering (AML) provisions apply to those providing services relating to transactions in cryptocurrencies. Key AML requirements include 'know your customer' (KYC) rules to verify identity—which in a crypto context, should enable linking of private keys with beneficial owners, at least in centralized exchanges—provide suspicious transaction reports (STRs), and attach customer information to transactions ('travel rules'). In the US, the applicability to transactions in crypto of AML rules was made clear in 2013 and, more widely, in 2015 the Financial Action Task Force (FATF)[75] issued guidance on the application of established guidelines. At EU level, prior regulations applied only to 'banknotes and coins, scriptural money and electronic money'[76] and so excluded cryptocurrencies; at the time of writing, a proposal issued by the Commission for an appropriately updated regulation, in line with FATF guidance, awaits Council approval. It was KYC provisions that provided the informational basis upon which the IRS has served 'John Doe' notices on crypto brokers seeking information on US taxpayers transacting USD 20,000 or more in cryptocurrency between 2016 and 2021.[77] In the UK, they enabled HMRC to write targeted letters to crypto owners reminding and informing them of their obligations.[78] Moreover, where 'tax crimes' are recognized as a predicate offence for money laundering, tax authorities in principle have direct access to the information collected by financial institutions under AML rules.

In practice, however, AML rules alone are commonly inadequate from a tax perspective (not only for crypto, but more generally). Not all jurisdictions fully comply with FATF guidelines,[79] and even where they do tax administrations may face obstacles in accessing the information they generate: surveying 28 of its members, OECD (2015) reports that only 20 per cent of tax administrations had direct access to STRs, leaving a heavy reliance on financial intelligence units to spontaneously share information that they deemed potentially tax-relevant. Nor is it in any case clear that even serious tax evasion will trigger STRs. Tax administrations' access to AML information may have improved since the OECD survey.[80] But extending AML rules to crypto transactions, important though it is for other purposes, is evidently insufficient for enabling their effective taxation. KYC rules might enable the authorities to know, for instance, that some individual cashed out a certain amount of cryptocurrency; but from the transactions prior to that recorded on the blockchain it will not be possible, without further information, to identify any associated capital gain or loss.[81]

Going beyond AML, the natural aspiration for tax administrations is to secure direct and automatic sharing with them of information on crypto transactions similar to that already quite widely in place for traditional financial

[74] See for example Baronchelli *et al*. (2022).
[75] The FATF is the international standard setter in AML and countering the financing of terrorism.
[76] See EC (2021).
[77] Department of Justice (2022*a*).
[78] As described, for example, in Saffery (https://www.saffery.com/insights/news/hmrc-letters-about-cryptocurrency-and-capital-gains-tax/).
[79] As of June 2022, FATF reports more than 20 jurisdictions as under 'increased monitoring' (FATF, 2022).
[80] OECD (2015) reports 57 per cent of surveyed jurisdictions as planning to ease their tax administrations' access to STRs.
[81] Innovation around the implementation of KYC rules can also be expected, as with the use of ATMs to acquire or dispose of cryptocurrencies using credit cards or, leaving even fewer traces, cash. In the UK, the perceived threat is such that these are illegal; in the US, while they are in principle subject to KYC rules implementation appears to be problematic (CNBC, 2021).

transactions. This is now the focus of considerable attention. In the US, the Infrastructure Improvement and Jobs Act (IIJA) of November 2021 includes two relevant provisions, requiring that: (i) a broadly defined set of digital service providers, potentially including even miners, report details of their customers' transactions to the IRS annually, just as with bonds and shares; (ii) all businesses report transactions in crypto assets of over USD 10,000, mimicking the pre-existing rule for cash payments. Both provisions take effect from tax year 2023. Similar measures are being adopted elsewhere. In Brazil, for instance, the Federal Revenue Service introduced regulations in 2019 requiring legal entities and individuals to report operations carried out with crypto assets.[82]

The application of reporting rules to domestic institutions may, however, drive transactions to either decentralized exchanges—generally not subject to KYC rules—or foreign ones that do not have, or in any case do not provide to the domestic authorities, information on holders. Cong *et al.* (2022) see signs of the latter effect at work in the US, finding that while IRS actions aimed at US taxpayers in general[83] appeared to increase legal avoidance activity in crypto markets, actions targeted at a specific US exchange (Coinbase) decreased activity on US exchanges.[84] In the UK, 22 per cent of survey respondents in HMRC (2022*a*) indicated that they preferred using foreign exchanges.

The key tool to address this is cross-border exchange of information, with the reporting of assets and income to the country of residence. Existing frameworks for this, however, were not constructed with cryptocurrencies, or crypto assets, in mind, so that they risk falling into a grey area.[85] Recognizing the problem, OECD (2022) sets out a framework for cross-border exchange between tax authorities of information on crypto transactions to parallel that already being applied for traditional financial assets and fiat currencies. This presumes that domestic authorities have in place requirements similar to those above, including for non-residents——so that implementation is likely some way off——but nonetheless provides a setting within which such rules can be developed with a substantial degree of commonality; the EU, for instance, appears to have been awaiting the outcome of this OECD work in order to build upon it among the member states. Importantly, the OECD framework reflects concern with the possible implications of crypto assets not only for income taxation but also for the VAT and sales taxes more generally: it includes a provision for the reporting of purchases by crypto assets of goods and services exceeding USD 50,000. General adoption of provisions along the lines of OECD (2022) remains, however, a somewhat distant prospect. In the (potentially long) meantime, all that entities will need do to avoid information reporting is 'to have their servers in a country where the authorities are willing to tolerate their existence'.[86] And, indeed, the experience has been that as some jurisdictions adopt international standards of information exchange, so activity tends to shift to others.[87]

Effective mechanisms of third-party reporting by centralized institutions involved in trading cryptocurrencies, including across borders, thus remain some way in the future. But it is at least possible to see how they might work. It is ultimately the use of decentralized exchanges and direct peer-to-peer transactions that pose the toughest problems.

Decentralized exchanges simply could not meet the kind of reporting obligations now imposed in the US. Perhaps for that reason, while not excluded from the reporting requirements of the Infrastructure and Jobs Act, they are also not explicitly included.[88] The reporting envisaged by OECD (2022) would include any decentralized exchange 'to the extent it exercises control or sufficient influence over the platform, allowing it to comply with the due diligence and reporting obligations'[89]—which, in most circumstances, it is not clear that they could do.

Whatever their current importance, however, the concern must be that as information reporting becomes more effective in relation to centralized trades, so activity will shift to decentralized forms upon which such rules cannot realistically be imposed. For these, more innovative methods may be needed.

One alternative approach, for example, may be to focus on requiring information reporting not (only) by things that resemble financial institutions but by the miners themselves.[90] These are involved in every cryptocurrency

[82] Individuals or legal entities conducting virtual currency operations for amounts exceeding 30,000 Reais (around USD 5,500) on a monthly basis must report this information. Cryptocurrency exchanges domiciled in Brazil for tax purposes must also provide information, annually, in relation to each user of their services.

[83] Reminding them of their obligations, and announcing a compliance programme focused on crypto.

[84] Fear of driving trading abroad may be one reason why the reporting requirements of the IIJA do not apply to non-US taxpayers.

[85] In the US, for instance, there is no explicit guidance on which foreign entities active in the crypto area are required to report under the Foreign Account Tax Compliance Act (FATCA) requirements.

[86] Makarov and Schoar (2021).

[87] As shown, for example, in Johannesen (2014).

[88] Decentralized exchanges were explicitly included in early versions of the Infrastructure and Jobs Bill, but in the final version the obligations apply to those 'regularly providing any service effectuating transfers of digital assets on behalf of another person' (Sec 8606 (a)(2)(D)). Whether 'effectuating' is to include the facilitation services provided by decentralized exchanges will perhaps be clarified in the accompanying regulations, which are still awaited. On language in the priori draft, see Forbes (https://www.forbes.com/sites/jasonbrett/2021/08/02/new-language-for-crypto-tax-reporting-excludes-decentralized-exchanges-miners-still-vulnerable/?sh=50ef00b55f56).

[89] OECD (2022, para 27, p. 51).

[90] This is suggested, from a perspective wider than that of taxation, by Makarov and Schoar (2022).

transaction (on-chain, at least) and, moreover—always useful for tax administration (though less welcome for the security of cryptocurrencies)[91]—are relatively few in number: Makarov and Schoar (2021) estimate that around 55–60 controlled more than half of all Bitcoin mining capacity at the end of 2020. In these respects, they are attractive points at which to require information reporting. One might even envisage going further and imposing at that point some corrective transaction tax (if determined to be helpful) and/or applying a (creditable) withholding obligation: a charge on each transaction paid to, and credited/refunded by, the relevant tax authority. Compliance with these obligations might be pursued in a number of ways. The traces that miners leave in the blockchain might help identify them as real entities. And/or incentives might be found to encourage their compliance, if not in the form of some explicit subsidy to the compliant (the optics of which might not be good), then in the familiar form of allowing some delay in remitting tax withheld to the authorities.[92]

For at least the immediate future, however, tax administrations in most countries must deal with potentially very limited directly usable information on cryptocurrency holdings and transactions. This of course does not mean they are helpless.

Working in their favour is the vast amount of information publicly available in permissionless blockchains. This provides scope for the application of techniques that have been developed for the forensic analysis of blockchain structures; some taste of this is provided by analyses discussed above, and there are, for example, firms that hold out the prospect of linking legal names, account numbers, and IP addresses to virtual asset service providers (such as crypto exchanges).[93] Artificial intelligence applications can draw on past experience to identify potentially tax-relevant behaviours, and there is of course room too for traditional investigative methods that seek links with information obtained outside the chain.

There also remain other and more standard measures to encourage self-reporting, such as taxpayer education, both general and targeted, and nudges. Large and successful actions can be used to send appropriately chilling messages: that the FBI, for example, 'is able to uncover the source of even the most sophisticated schemes and bring justice to those who try to exploit the security of our financial infrastructure', and that a seizure of NFTs in the UK 'serves as a warning to anyone who thinks they can use crypto assets to hide money from HMRC'.[94] Given the profundity of the fundamental challenges posed by pseudonymity, the astonishingly rapid pace of change in the area, its technical complexity, the vast information gaps, the uncertainties ahead, and a sense that the tide has not yet turned in the battle to incorporate crypto properly into the wider tax system, one might detect in such statements some element of whistling in the dark. It is clear, in any event, that living up to these confident assertions is not easy.

(iii) Value added and sales taxes

Much of the discussion of the tax implications of cryptocurrencies, and of crypto assets more generally, has focused, implicitly or otherwise, on the taxation of income, and especially of capital gains. Looking forward, however, some of the greatest risks to the broader tax system may be those arising in relation to the VAT and sales taxes.

The use of cryptocurrencies to acquire goods and services directly is apparently modest now, and not a feature of everyday life (even where Bitcoin is legal tender). And indeed in some respect, as also seen above, current tax rules may impede its use as a means of payment. If such use were to become widespread, however, it could create potentially significant dangers to the integrity of VAT and sales tax systems. One obvious risk is that widespread use of cryptocurrencies could facilitate the underreporting of final sales. Qualitatively, this is not a new problem: indeed, a focus of tax administration for decades has been to counter this risk, especially in relation to cash purchases. And that has been done with some success. Crypto, however, may open up a new front in this battle, waged with new and complex weapons.

The first line of defence is imposition of a legal requirement on all businesses to report large crypto transactions, of the kind, noted above, that is now in place in the US and envisaged by OECD (2022). Such rules are clearly not sufficient to bring the dangers to tax systems set out above under control: they are neither self-enforcing nor all-encompassing. And they face the same difficulties as above from decentralized trading or use of non-reporting

[91] This is because concentration in mining increases the risk of a '51% attack', by which malicious actors become able to manipulate the contents of the blockchain.

[92] One could also conceive of levying some form of penalty on entities found to be using non-compliant miners. This though would require that users be able to select miners *ex ante*, an option not now available and which, by limiting the number of miners, might again have implications for the security of the blockchain itself.

[93] As for instance at https://ciphertrace.com/solutions/.

[94] Quotations respectively from FBI Deputy Director (Department of Justice, 2022*b*) and HMRC Deputy Director of Economic Crime (Guardian, 2022). As noted, the UK seizure was not a specifically crypto-related tax offence but was related to the financing of a potential VAT fraud. HMRC has also let it be known that it 'is opening increasing numbers of investigations into people who have only disclosed modest amounts on their tax returns whilst transacting major amounts in crypto. HMRC is said to have [about] 20 new criminal investigations underway involving crypto and for every criminal investigation there are likely to be dozens of civil investigations' (Park, 2022).

foreign exchanges. But they are close to necessary, in generating red flags, clues for audit, and increasing the downside risks of evasion.

For the VAT, further challenges may arise in the use of cryptocurrencies as a convenient device for fraud, for example in creating carousels that enable refunds to be claimed for tax that has not been paid.[95] This is again not a new problem, but one that the use of cryptocurrencies may cast in a new form.

Little systematic thought appears yet to have been given to protecting sales taxation and, bringing additional complexity, the VAT, against these challenges. The risks, for now, appear more latent than real. But this can change.

VI. Conclusion

The future of cryptocurrencies is highly uncertain. For some, they are a bubble that (happily mixing metaphors) will sooner or later crash and burn. To others, they will prove the foundation for fundamental innovations in decentralized finance. In either case, however, tax systems need to accommodate them with a coherence, clarity, and effectiveness that, having been constructed without crypto assets in mind, they currently lack. They need to do so, moreover, in the context of continuing rapid and complex innovation, on the basis of limited information, and while balancing the core objectives of securing efficiency, fairness, and revenue in taxation against the risk of stifling innovation. The challenges are both conceptual and, still more, practical.

Conceptually, the dual nature of cryptocurrencies as both investment assets and means of payment—the latter, though less prominent than the former, being a primary purpose for their development—creates potential difficulty in capturing capital gains and losses in their role as assets without thereby constructing obstacles to their use as currency. For the VAT and sales taxes, while many issues of detail arise, the critical step is to ensure that cryptocurrency is treated the same way as national currencies. What is needed for income tax purposes (for instance providing exemptions for reasonable personal use as currency) will depend on existing national structures for the treatment of gains and losses and of foreign currency transactions.

Questions also arise as to a potential corrective role for taxation, whether, for example, to address internalities associated with gambling or as a stop-gap to dampen change pending the implementation of effective regulation. Much more clear cut, and now becoming the focus of some action, is the case for some charge—ideally as a part of a wider carbon tax, or if not as a sector-specific charge—to address the significant climate impact of proof-by-work consensus mechanisms.

It is in implementation, however, that cryptocurrencies pose the most severe problems. This is because their essence, and the core motivation behind their development, is precisely to avoid placing trust in centralized institutions of a kind that might be able to provide information to the tax authorities, or perhaps to levy some kind of withholding tax. The first step for governments, nonetheless, is to apply AML rules and third-party reporting requirements where they can, as the US has recently done. The risk, however, is that transactions will to some degree migrate to forms (on decentralized exchanges or directly peer-to-peer) that no third party even sees. It might be, however—somewhat ironically, in terms of the original vision for crypto—that investors come to place more trust in well-regulated centralized institutions than in the 'Wild West' of decentralized trading. More speculatively, miners—who do see every on-chain transaction in non-stablecoins—might in principle be given a role in tax reporting/withholding, consistent with the general principle of tax administration favouring collection, where possible, at a relatively small number of upstream points. These difficulties are amplified, moreover, by the ease with which crypto transactions are conducted across borders, so that domestic tax measures might also result in trades shifting to non-reporting platforms abroad. The OECD (2022) has developed a framework for extending current arrangements for cross-border information to crypto, but implementation remains some way in the future and in any case will not in itself resolve the challenges posed by decentralized trading.

It is hard to assess the extent to which the pseudonymity of crypto facilitates tax evasion (or indeed is associated with outright criminality, though that seems to account for a declining share of activity in cryptocurrencies). There is, however, evidence of a degree of compliance, at least for the US, with about 1 per cent of returns reporting crypto sales and signs of significant avoidance activity. In terms of the amounts at stake, rough calculations suggest that in 2021, a 'good' year for cryptocurrencies, a global tax at 20 per cent on accrued capital gains might have raised around USD 300 billion (which is about 12 per cent of global revenue from the corporate income tax). But in the 'bad' year of 2022, revenue would likely be eroded by large capital losses.

As for those likely to have gains subject to tax, while information is limited, there is strong evidence that crypto wealth is highly concentrated, even more so than ownership of equities. This, it seems, is not just a matter of a

[95] On the nature of carousel fraud, see, for example, Keen and Smith (2006).

handful of crypto billionaires, or of the few who have become wealthy by investing in crypto, but of holdings by a wider but unknown set of the rich whose wealth derives from other sources. It may be, of course, that the deepest problems here are not distinctive to crypto but apply to the capital gains taxation of other assets too, especially at the top of the income distribution. Worth bearing in mind, however, is that, at least in the UK and US, many holders of cryptocurrencies are far from being among the richest, with incomes that are lower than other investors and often no more than moderate.

Such literature as there is on the taxation of cryptocurrencies has focused on income tax aspects, which is indeed that most relevant to the taxation of the rich that is the topic of this issue. Much less attention has been paid to the implications for sales taxation and, especially, the VAT. These may, however, be profound, as cryptocurrencies potentially create problems similar to those associated with the use of cash, with which VATs have long struggled, and perhaps create new opportunities for fraud. It is in this area that the risks to existing tax systems created by the use of cryptocurrencies may prove most profound—not least in emerging and developing economies in which demand for crypto appears relatively strong while tax administration is relatively weak.

Appendix: the mechanics of cryptocurrencies[96]

A *distributed ledger technology* (DLT) is one that makes some database available, for inspection and/or amendment, to authorized users, with a protocol of consensus in place to ensure that—without any need for a central authority—all entries are accurate and protected against tampering. The best known and most widely used DLT is blockchain, the distinctive feature of which is that transactions are added to the database sequentially, with approval of each new block entailing confirmation of previous blocks, and the use of encryption to make it extremely difficult to change earlier entries. For many applications, such as tracking the movement of goods, the system is ‘permissioned’, meaning that access and/or rights are in some way limited, and authority to introduce and confirm changes restricted to particular users. Cryptocurrencies, however, are generally *permissionless*, meaning that access to the database is fully public; this requires particular measures to ensure trust in the database and avoid use of the same funds more than once: ‘double-spending’.

Protection against tampering in cryptocurrencies rests on cryptographic methods that enable private information to be encrypted in such a way that its accuracy can be verified without revealing that information itself.[97] To implement this, users have a *private key* (or ‘address’) that is encrypted into a *public key* that is known to all users, but from which it cannot be inferred. Details of a proposed transaction, along with the public key, are then broadcast to all participants. The accuracy of that information and availability of the necessary coins is easily checked. The mechanism for validation, however—meaning addition to the chain—is made costly in order to deter tampering and prevent double-spending, which in turn requires some reward for doing so. Under *proof of work*, this is done by having validators (‘miners’) compete to solve—more accurately, guess a solution to—a complex numerical problem, requiring extensive computing power, in return for which they receive an allocation of the crypto currency and/or a fee. Under *proof of stake*, the task of verification is allocated probabilistically in proportion to an amount of crypto that is staked; the reward for such validation is again some amount of the cryptocurrency, with a loss of stake in the event of failure or misrepresentation. Once confirmed in this way, the new block is added to the chain.

Stablecoins, being backed or with their supply controlled algorithmically, do not require verification in this way.

Private keys, from which public addresses are derived by encryption, are held in electronic/digital *wallets*, which may be held offline (‘cold’) or by service providers, whether as custodians (taking control of the key, executing trades at the customer’s request) or simply providing security.

References

Abadi, J., and Brunnermeier, M. (2022), ‘Blockchain Economics’, Federal Reserve Bank of Philadelphia WP 22-15.

Allingham, M. G., and Sandmo, A. (1972), ‘Income Tax Evasion: A Theoretical Analysis’, *Journal of Public Economics*, **1**(3), 323–38.

Alnasaa, M., Gueorguiev, N., Honda, J., Imamoglu, E., Mauro, P., Primus, K., and Rozhkov, D. (2022), ‘Crypto-assets, Corruption, and Capital Controls: Cross-Country Correlations’, *Economics Letters*, **215**(100), 110492.

[96] Many subtleties are skimmed over here; for more detail, see, for instance, Hallaburda *et al*. (2022) and Box 8-2 of Council of Economic Advisers (2023).

[97] This rests on the use of non-invertible ‘hash’ functions which map a number or text of arbitrary length into a unique number from which it cannot be inferred.

Alstadsæter, A., Johannesen, N., and Zucman, G. (2019), 'Tax Evasion and Inequality', *American Economic Review*, **109**(6), 2073–103.

Alvarez, F. E., Argente, D., and Van Patten, D. (2022), 'Are Cryptocurrencies Currencies? Bitcoin as Legal Tender in El Salvador', NBER WP 29968.

Auer, R., and Tercero-Lucas, D. (2022), 'Distrust or Speculation? The Socioeconomic Drivers of US Cryptocurrency Investment', *Journal of Financial Stability*, **66**, 101066.

— Cornelli, G., Doerr, S., Frost, J., and Gambacorta, L. (2022), 'Crypto Trading and Bitcoin Prices: Evidence from a New Database of Retail Adoption', BIS Working Papers No. 1049.

Avi-Yonah, R., and Mohanad, S. (2022), 'A New Framework for Taxing Cryptocurrencies', University of Michigan Public Law Research Paper No. 22-014.

Baronchelli, A., Halaburda, H., and Teytelboym, A. (2022), 'Central Bank Digital Currencies Risk Becoming a Digital Leviathan', *Nature Human Behaviour*, **6**, 907–9.

Beck, T., Janfils, M., and Kpodar, K. (2022), 'What Explains Remittance Fees? Panel Evidence', IMF WP No 63/2022.

Bloomberg (2021), 'Sotheby's Makes $100 Million in NFT Sales With Younger Audience', https://www.bloomberg.com/news/articles/2021-12-15/sotheby-s-makes-100-million-in-nft-sales-with-younger-audience#xj4y7vzkg

Board of Governors (2022), 'Economic Well-being of US Households in 2021', https://www.federalreserve.gov/publications/report-economic-well-being-us-households.htm

CBS (2021), 'There May Now be as Many as 100,000 Bitcoin Millionaires', 23 February, https://www.cbsnews.com/news/bitcoin-millionaires-100k/

Chainalysis (2022*a*), 'The 2022 Crypto Crime Report', https://go.chainalysis.com/2022-Crypto-Crime-Report.html

— (2022*b*), 'The 2022 Geography of Cryptocurrency Report', https://go.chainalysis.com/geography-of-crypto-2022-report.html

— (2022*c*), 'The Chainalysis State of Web3 Report', https://go.chainalysis.com/2022-web3-report.html.

Clotfelter, C. T. (2005), 'Gambling Taxes', ch. 4 in *Theory and Practice of Excise Taxation: Smoking, Drinking, Gambling, Polluting, and Driving*, Oxford, Oxford University Press.

CNBC (2018), 'Warren Buffett: Cryptocurrency Will Come to a Bad Ending' [Interview with Warren Buffett on 10 January 2018], https://www.cnbc.com/video/2018/01/10/warren-buffett-cryptocurrency-will-come-to-a-bad-ending.html

— (2021), 'Cash In, Fraud Out: Criminals Target Bitcoin ATMs as Crypto Popularity Surges', https://www.cnbc.com/2021/11/09/bitcoin-atms-criminals-target-cryptocurrency-transactions.html.

Cnossen, S., and Jacobs, B. (2022), 'Problemen met een partiële vermogensaanwasbelasting', Free University of Amsterdam.

Cong, L. W., Landsman, W., Maydew, E., and Rabetti, D. (2022), 'Tax-loss Harvesting with Cryptocurrencies', mimeo, Cornell University.

Council of Economic Advisers (2023), 'Economic Report of the President'.

Department of Justice (2022*a*), 'Court Authorizes Services of John Doe Summons Seeking the Identities of US Taxpayers Who Have Used Cryptocurrency', press release, 16 August.

— (2022*b*), 'Two Arrested for Alleged Conspiracy to Launder $4.5 Billion in Stolen Cryptocurrency', press release, 8 February.

EC (2011), 'Financial Transactions Tax: Making the Financial Sector Pay its Fair Share', Brussels, European Commission, https://ec.europa.eu/newsroom/growth/items/45441/.

— (2021), 'Proposal for a Regulation of the European Parliament and of the Council on Information Accompanying Transfer of Funds and Certain Crypto-assets (Recast)', European Commission, COM(2021) 411 final 2021/0241 (COD), Brussels, 20 July.

Europol (2021), 'Cryptocurrencies: Tracing the Evolution of Criminal Finances', Europol Spotlight, Luxembourg, Publications Office of the European Union.

FATF (2022), 'Jurisdictions under Increased Monitoring', The Financial Action Task Force, https://www.gov.uk/government/consultations/call-for-evidence-the-taxation-of-decentralised-finance-involving-the-lending-and-staking-of-cryptoassets

Foley, S., Karlsen, J. R., and Putniņš, T. J. (2019), 'Sex, Drugs, and Bitcoin: How Much Illegal Activity is Financed through Cryptocurrencies?', *Review of Financial Studies*, **32**(5), 1798–853.

Forbes (2021), 'The Cryptocurrency Tycoons on Forbes' 2021 Billionaires List', https://www.forbes.com/sites/johnhyatt/2021/04/06/the-cryptocurrency-tycoons-on-forbes-2021-billionaires-list/?sh=1c0e59cf25e2

— (2022), 'The Richest Crypto and Blockchain Billionaires in the World 2022', https://www.forbes.com/sites/johnhyatt/2022/04/05/the-richest-crypto-and-blockchain-billionaires-in-the-world-2022/?sh=17debf4b580d

Guardian (2022), 'HMRC Seizes NFTs for First Time amid Fraud Inquiry', https://www.theguardian.com/technology/2022/feb/14/hmrc-seizes-nfts-for-first-time-amid-fraud-inquiry

Hallaburda, H., Gans, J., and Gandal, N. (2022), 'The Microeconomics of Cryptocurrencies', *Journal of Economic Literature*, **60**, 971–1013.

Hebous, S., and Vernon, N. (forthcoming), 'Cryptocarbon: How much is the Corrective Tax?', IMF WP, Washington, DC, International Monetary Fund.

HMRC (2022*a*), 'Individuals Holding Cryptoassets: Uptake and Understanding', London, HMRC.

— (2022*b*), 'The Taxation of Decentralised Finance Involving the Lending and Staking of Cryptoassets—Call for Evidence', London, HMRC, https://www.gov.uk/government/consultations/call-for-evidence-the-taxation-of-decentralised-finance-involving-the-lending-and-staking-of-cryptoassets

Hoopes, J. L., Menzer, T. S., and Wilde, J. H. (2022), 'Who Sells Cryptocurrency?', mimeo, University of North Carolina.

IMF (2023), 'Elements of Effective Policies for Crypto Assets', Policy Paper, Washington, DC, International Monetary Fund.

InsiderIntelligence (2022), '34 Million US Adults Own Cryptocurrency', April, https://www.insiderintelligence.com/insights/us-adults-cryptocurrency-ownership-stats/

ITR (2022), 'This Week in Tax: Crypto Exchanges Fall Foul of India's GST Law', *International Tax Review*, 10 April, https://www.internationaltaxreview.com/article/2a6ab6uyjg7ixq6atojy8/this-week-in-tax-crypto-exchanges-fall-foul-of-indias-gst-law

Johannesen, N. (2014), 'Tax Evasion and Swiss Bank Deposits', *Journal of Public Economics*, **111**, 46–62.

Joint Committee on Taxation (2021), 'Estimated Revenue Effects of the Provisions in Division H of an Amendment in the Nature of a Substitute to HR 3684', 2 August, JCX-33-21.

Keen, M., and Smith, S. (2006), 'VAT Fraud and Evasion: What Do We Know and What Can Be Done?', *National Tax Journal*, **59**, 861–87.

Makarov, I., and Schoar, A. (2021), 'Blockchain Analysis of the Bitcoin Market', NBER WP 29396.

— — (2022), 'Cryptocurrencies and Decentralized Finance (DEFi)', BIS Working Papers No. 1061.

Matheson, T. (2012), 'Security Transaction Taxes: Issues and Evidence', *International Tax and Public Finance*, **19**, 884–912.

MotleyFool (2021), 'Tim Cook Owns Cryptocurrency; And So Do 68% of American Millionaires', 23 November, https://www.fool.com/research/american-millionaire-crypto-investors/

— (2022), 'Study: Over 46 Million Americans Likely to Buy Crypto in the Next Year', 21 June, https://www.fool.com/the-ascent/research/study-americans-cryptocurrency/

Nguyen, X.-T., and Maine, J. A. (2023), 'Crypto Losses', mimeo, University of Washington Law School.

OECD (2015), 'Improving Co-operation Between Tax and Anti-money Laundering Authorities', Paris, Organization for Economic Cooperation and Development.

— (2020), 'Taxing Virtual Currencies An Overview of Tax Treatments and Emerging Tax Policy Issues', Paris, Organization for Economic Cooperation and Development.

— (2022), 'Crypto-asset Reporting Framework and Amendments to the Common Reporting Standard', Paris, Organization for Economic Cooperation and Development.

Panetta, F. (2023), 'Caveat Emptor Does Not Apply to Crypto', *Financial Times*, 4 January.

Park, A. (2022), 'Going Digital—HMRC Begins Seizing Crypto Assets', https://www.paminsight.com/epc/article/going-digital-hmrc-begins-seizing-crypto-assets#

PwC (2021), 'PwC Annual Global Crypto Tax Report', https://www.pwc.com/kz/en/crypto-currency.html

Sarfo, N. A. (2022), 'Central Bank Digital Currencies Raise Tax Questions', *Tax Notes International*, **107**, 989–92.

Tasca, P., Liu, S., and Hayes, A. (2022), 'The Evolution of the Bitcoin Economy: Extracting and Analyzing the Network of Payment Relationships', *Journal of Risk Finance*, **19**(2).

Thiemann, A. (2021), 'Cryptocurrencies: An Empirical View from a Tax Perspective', JRC Working Papers on Taxation and Structural Reforms No 12/2021.

Waerzeggers, C., Aw, I., and Cheng, J. (2023), 'Taxing Stablecoins', International Monetary Fund Fintech Note 2023/002.

Wiseman, S. A. (2016), 'Property or Currency? The Tax Dilemma Behind Bitcoin', *Utah Law Review*, **2**, 417–40.

World Bank (2018), 'Cryptocurrencies and Blockchain', Washington, DC, World Bank.

Zucman, G. (2013), 'The Missing Wealth of Nations: Are Europe and the US Net Debtors or Net Creditors?', *The Quarterly Journal of Economics*, **128**(3), 1321–64.