INTERNATIONAL MONETARY FUND

# Decrypting Crypto: How to Estimate International Stablecoin Flows

Marco Reuter

**WP/25/141**



**2025**
JUL

WORKING PAPER

 WP/25/141

**IMF Working Paper**
Research Department

**Decrypting Crypto: How to Estimate International Stablecoin Flows**
**Prepared by Marco Reuter***

Authorized for distribution by Maria Soledad Martinez Peria
June 2025



**ABSTRACT:** This paper presents a novel methodology—leveraging a combination of AI and machine learning to estimate the geographic distribution of international stablecoin flows, overcoming the “anonymity” of crypto assets. Analyzing 2024 stablecoin transactions totaling $2 trillion, our findings show: (i) stablecoin flows are highest in North America ($633bn) and in Asia and Pacific ($519bn). (ii) Relative to GDP, they are most significant in Latin America and the Caribbean (7.7%), and in Africa and the Middle East (6.7%). (iii) North America exhibits net outflows of stablecoins, with evidence suggesting these flows meet global dollar demand, increasing during periods of dollar appreciation against other currencies. Further, we show that the 2023 banking crisis significantly impeded stablecoin flows originating from North America; and finally, offer a comprehensive comparison of our data to the Chainalysis dataset.


| | |
|---|---|
| JEL Classification Numbers: | C58, F31, G15 |
| Keywords: | stablecoins; capital flows; capital flight; capital flow management measures (CFMs); crypto assets; currency substitution; dollar demand |
| Author’s E-Mail Address: | mreuter@imf.org |


* I thank Cage Englander for excellent research assistance. I am grateful to Itai Agur, Alexander Copestake, Alessia De Stefani, Andrés Fernández, and Maria Soledad Martinez Peria for comments; to Gordon Liao, and Ulf Lewrick for excellent discussions; and to audiences at seminars at the IMF, the IMF’s 7th An-nual Macro-Financial Conference, and the SNB-CIF Conference on Cryptoassets and Financial Innovation.

WORKING PAPERS

# Decrypting Crypto: How to Estimate International Stablecoin Flows

Prepared by Marco Reuter

# 1 Introduction

Policymakers are increasingly wary of the popularity of crypto assets and have called for better monitoring of crypto transactions and international crypto asset flows (BIS (2023), EU (2023), G7 (2023) , FATF (2023), FSB (2023), IMF (2023), US Treasury (2023)). At the same time, recent research shows that crypto assets are increasingly used for international transactions, particularly when capital flow measures make it difficult to use traditional channels (von Luckner et al. (2023), von Luckner et al. (2024)), and that they could potentially be sizable (Cardozo et al. (2024), Cerutti et al. (2024), Auer et al. (2025)). However, estimating international crypto asset flows remains challenging due to the opaque nature of crypto assets.

The main contribution of this paper is the development of a novel method that enables the identification of the geographic regional origin of crypto wallets, facilitating the measurement of international stablecoin flows.[1] Before detailing our method, we address a common misconception—contrary to popular belief, the vast majority of crypto assets do **not** provide anonymity. Every transaction is publicly recorded on a freely accessible ledger known as a blockchain. The perception of anonymity arises because blockchain data is pseudonymized; rather than recording personal information such as names or residences, blockchains log only the wallet addresses of senders and receivers. A wallet address, typically a long hexadecimal string such as
`'0xdFDEe1155E1dd7c01774560C6E98C41B7da945dB'`, does not directly reveal personal information about the user. The key challenge in mapping the geography of crypto asset flows is supplementing blockchain data with useful information about senders and receivers. Our methodology addresses this challenge by enabling the estimation of the geographic region of *any arbitrary* self-custodial wallet[2] in the Ethereum ecosystem.

To estimate the geographic region of self-custodial wallets (we assign wallets to one of the following five regions: Africa and the Middle East, Asia and the Pacific, Europe, North America, and Latin America and the Caribbean), our methodology involves obtaining geographic information for a subset of wallets through two distinct approaches. First, we leverage domain names assigned to wallets through systems such as the Ethereum Name System (ENS).[3] We employ a large language model (LLM) to infer linguistic and cultural

[1]A stablecoin is a crypto asset with its value pegged to a fiat currency, most often the US dollar. The most popular stablecoins are Tether's USDT and Circle's USDC, boasting a combined market capitalization exceeding $215 billion (June 2025).

[2]A self-custodial wallet is a type of crypto wallet where the user has full control and responsibility over their funds, without relying on third-party intermediaries for custody.

[3]Name systems allow users to replace the long hexadecimal strings with human-readable names. A similar system, the *Domain Name System (DNS)* is at the core of the internet, replacing numerical IP-addresses

markers—such as language, script, or regional references—that suggest a wallet's likely region. Second, we identify wallets that frequently transact with centralized exchanges (CEXs) targeting specific regional markets, assuming that a wallet predominantly interacting with, for example, a Latin America focused exchange is likely from that region. These two methods provide an ad hoc regional classification for a subset of wallets, which we then use as labeled training data to train a machine learning model for classification of arbitrary wallets.

The core of our approach lies in leveraging this training data to train a machine learning model to recognize patterns in on-chain activity that are indicative of a wallet's geographic origin. We construct features capturing wallets' behavioral and transactional characteristics, including time-of-day activity patterns, adherence to daylight savings time, interactions with certain centralized exchanges, and engagement with popular ERC-20 tokens and smart contracts. By learning region-specific patterns the trained model can estimate the geographic region of any arbitrary self-custodial wallet. The identifying assumption of the methodology is, that conditional on the features we selected to train the model, wallets that are in the training set and those outside the training set exhibit the same patterns. The methodology then enables us to map the geographic distribution of wallets, which we can then leverage to map international stablecoin flows.

Using this approach, we analyze almost 6 million domain names and billions of on-chain transactions to construct the training dataset. Some examples of regions assigned by analyzing domains names with the use of a LLM are, "pijiu" (Chinese for "beer") which is assigned to Asia and the Pacific, and الكيماويات (Arabic for "chemicals") which is assigned to Africa and Middle East. We validate these ad hoc classifications through time-of-day based activity profile analysis. By adjusting wallet transaction timestamps to regional time zones (e.g., UTC+8 for Asia and the Pacific, UTC-6 for North America), we observe distinct peaks in activity during local daytime hours (e.g., 10 AM to 10 PM) and lows during nighttime, supporting the accuracy of our regional assignments. Additionally, we document differences in activity patterns due to daylight savings time (DST) in regions like North America (where DST is common) and a lack of such differences in regions where DST is almost non-existent, like Asia and the Pacific.

We then train a Gradient Boosted Decision Tree model on this dataset, which achieves an overall accuracy of 65% in predicting geographic regions. For context, random guessing with five regions yields 20% accuracy. We then apply the model to predict the region of any arbitrary self-custodial wallet. Using these predictions, we map international stablecoin flows for 2024. Our analysis captures approximately 138 million transactions totaling $2,019 billion, with an average transaction size of $14,630.

with domain names in the common "www.xyz.com" format.

We document significant regional variation in stablecoin usage. In absolute terms, we estimate that Asia and the Pacific lead with the highest stablecoin activity (inflows: \$407bn, outflows: \$395bn, intraregional flows: \$209bn), followed by North America (inflows: \$363bn, outflows: \$417bn, intraregional flows: \$216bn). However, relative to GDP, Africa and the Middle East, and Latin America and the Caribbean stand out, with stablecoin usage reaching 6.7% and 7.7% of GDP, respectively. Additionally, intraregional flows in these two regions are notably lower, accounting for 14% and 12% of total flows originating from the region, compared to, for example, 34% in North America. This suggests that stablecoin use in Africa and Latin America is predominantly international, possibly driven by use cases such as remittances. Further, there is also significant heterogeneity in average transaction sizes, spanning from \$11,493 (Asia and Pacific) to \$35,016 (North America). Intuitively, regions with higher GDP/capita (e.g., North America and Europe) exhibit the largest average transaction sizes, while average transaction sizes in the other regions are significantly lower.

We also document regional heterogeneity in usage of different stablecoins (i.e., USDC vs USDT) and different CEXs. Tether's USDT is more popular in regions with more emerging economies—Africa and the Middle East, Asia and the Pacific, and Latin America and the Caribbean—while Circle's USDC is more prevalent in regions with more advanced economies, i.e., Europe and North America. Regarding crypto exchanges, we find that Binance is preferred in emerging market regions, whereas Coinbase leads in North America.

Calculating bilateral net flows highlights North America as the primary source of stablecoin outflows into all other regions of the world, which we estimate to amount to \$54bn in 2024.[4] Using this observation, we validate our dataset by providing evidence that net stablecoin flows from North America into other geographic regions increase when domestic currencies are weak, complementing a related relationship between stablecoin flows and high inflation established by Auer et al. (2025). Additionally—after training an additional model to provide country-level estimates for China—we show that a similar pattern holds for stablecoin flows into China. That is, they increase significantly when the US-Dollar appreciates vis-à-vis the Chinese Renminbi. This suggests that stablecoins could increasingly be serving as an instrument to meet global demand for dollars, particularly in regions where access to traditional dollar markets is constrained (see e.g., Calvo and Reinhart (2002), Gopinath and Stein (2021) for global demand on dollars, and Aramonte et al. (2022) for evidence on stablecoin demand in emerging markets).

We further validate our dataset—and establish novel empirical evidence of the connection between stablecoins and the banking system—by documenting how stablecoin flows were

[4]These flows are likely accompanied by countervailing flows in fiat currency or goods into North America, which we do not observe.

disrupted during the March 2023 banking crisis. Many of the banks affected in the crisis provided banking services to CEXs and stablecoin issuers, which are crucial for settling the fiat currency leg in stablecoin issuance and redemption. In a differences-in-differences regression, with North America as the treated region and the other regions a control, we show that the crisis significantly reduced stablecoin flows originating from North America.

Finally, we also offer the first comprehensive dataset that can be compared to the commercially available Chainalysis dataset, that has been extensively used in the literature (e.g., in Cardozo et al. (2024), Cerutti et al. (2024), Auer et al. (2025)). For context, we briefly describe the methodology at the core of that dataset: Chainalysis infers the geographic origin and destination of stablecoin transactions by focusing on transactions involving CEXs. To assign regions, Chainalysis uses web traffic data associated with CEXs. That is, they obtain data about the number of users that access the corresponding websites of the CEX, broken down by country. They assume that the geographic distribution of a platform's website visitors reflects its user base. For instance, if a significant share of a CEX's web traffic originates from Brazil, Chainalysis attributes a proportional share of that exchange's stablecoin transactions to Brazilian users. For identification, they assume that users do not use VPNs[5], when accessing these websites, as VPNs are often used to misrepresent the true country of origin. Further, this approach assumes that transaction sizes are uniform across users, regardless of their country or region. When a stablecoin is sent from one CEX to another, the flow is then broken down according to web-traffic proportions and assigned as an in/outflow between the respective countries.

In contrast, our approach focuses on estimating the geographic region of self-custodial wallets. For identification, we assume that conditional on the features we selected in the machine learning model, wallets that are in the training set and those outside the training set exhibit the same patterns. Advantages of our approach are that we do not assume that users do not use VPNs or that users in different regions, on average, have the same transaction size. In fact, we provide evidence that invalidates this assumption. A disadvantage of our approach however, is the lower granularity of the estimates. While we can (largely) only provide region-level estimates, the Chainalysis dataset provides estimates on a country-level basis.

Comparing both datasets, we find both significant agreements and disagreements. For example, both datasets roughly agree on the volume of stablecoin flows in 2024, with our estimate being \$2,109bn, while Chainalysis estimates \$1,730bn. Differences in the over-

[5] A VPN, or Virtual Private Network, is a service that encrypts internet traffic and routes it through a remote server, often located in a different country, concealing the user's location and enhancing online privacy and security.

all quantities can be explained by differences in coverage of the underlying blockchains, and coverage of certain transactions. They further agree that Asia and Pacific, and North America exhibit significant flows in terms of absolute volumes, while Africa and the Middle East, and Latin America and Caribbean lead in flows relative to GDP. On net flows, both datasets broadly agree that stablecoins largely from from North America to the other regions. However, disagreement arises with the "indirect" category in the Chainalysis data, which estimates a reversal in the direction of flows, contradicting both our estimates and those of the "direct" category in the Chainalysis dataset.[6] Further, there is significant disagreement regarding the use of stablecoins in China. For 2024, we estimate 5.5 times more gross stablecoin flows involving China (i.e., \$153bn vs \$28bn) and 100 times more net flows of stablecoins into China (i.e., \$18bn vs \$0.18bn). We believe that the no VPN assumption of Chainalysis is systematically violated for China. For example, we estimate Binance to be the most significant CEX in China, and that Binance drives \$11 billion in net stablecoin inflows to Chinese self-custodial wallets, despite its website being inaccessible without a VPN.[7]

**Related Literature.** This paper contributes in particular to the literature that focuses on the measurement and drivers of international crypto flows, and in general to the literature that studies the connection between crypto assets on macroeconomic variables.

In terms of methodology, the closest paper is Athey et al. (2016). They estimate the region of self-custodial Bitcoin wallets in 2015 on a small scale by identifying the region of 2,858 addresses scraped from online forums and train a random forest model for classification, allowing them to generate empirical evidence in support of their theoretical model on Bitcoin adoption. Relative to their work, we contribute by (1) generating a much larger dataset ($\sim$350,000 wallets) by analyzing wallet domain names, which had yet to be invented in 2015; (2) leveraging this dataset to train a model that can be used to estimate the region of any arbitrary self-custodial wallet and (3) applying the estimates to comprehensively map international crypto asset flows. Other papers (Meiklejohn et al. (2013) and Makarov and Schoar (2021)), have used heuristic clustering and behavior-based classification to link wallets, but have not attempted to specifically make geographic predictions about wallets.

Recently, a literature that attempts to quantify international crypto flows and study their drivers has emerged (Cardozo et al. (2024), Cerutti et al. (2024), Auer et al. (2025)), that has leveraged the dataset by Chainalysis as the basis for their analysis. Relative to

[6] Chainalysis splits flows into *direct* flows, which flow directly from CEX to CEX, and *indirect* flows, which they attempt to trace while passing through self-custodial wallets en route between CEXs. For a more detailed explanation, see https://docs.markets.chainalysis.com/#flow-categories.

[7] Coincidentally, Binance hosts a blog post on how to use Binance via VPN from China (https://www.binance.com/en/square/post/17293683766001).

these papers, we contribute by introducing a novel methodology to estimate international crypto flows that does not rely on the Chainalysis dataset and its assumptions (no VPN usage, same transaction size in every country), by producing the first dataset that allows for a comprehensive comparison to the Chainalysis data, and by outlining some novel drivers of stablecoin flows, such as exchange rates and the March 2023 banking crisis.[8] Other papers, such as, von Luckner et al. (2023) and von Luckner et al. (2024) have focused on providing qualitative evidence that crypto assets are used for circumventing capital controls and to facilitate capital flight. Relative to these, we contribute by providing quantitative estimates. Further, von Luckner et al. (2023) had to exclude dollar denominated flows from their analysis of peer-to-peer transactions for identification, which we specifically focus on by analyzing stablecoins.

Our paper also connects to a wider literature, linking macroeconomic conditions and crypto usage more generally. For example, Cong et al. (2023) provide evidence that households adopt crypto assets and stablecoins when they expect higher inflation. Our data indirectly supports this finding by showing that stablecoin usage in regions with more emerging economies is relatively higher as a fraction of GDP. Alnasaa et al. (2022) show that crypto usage is positively correlated with perceived corruption and capital controls, while Arbalik et al. (2021) document correlation between Bitcoin volatility and capital flows. Focusing on China, Hu et al. (2021) find evidence of capital flight through Bitcoin in China, a finding that our country-specific analysis on China supports, extends to stablecoins, and is able to quantify.

The rest of the paper is structured as follows: section 2 provides some background information about international stablecoin flows, section 3 provides an overview of the data used, section 4 explains the methodology used, section 5 explains the results of the classification model, section 6 maps international stablecoin flows, section 7 analyzes economic drivers of stablecoin flows, and section 8 compares our dataset with the Chainalysis dataset.

## 2 Background

In this section, we offer background information about how international crypto flows operate in practice. For that, we break down the typical transaction chain, identifying which legs of the transaction are observable in our data. Figure 1 provides a schematic overview that summarizes how international crypto flows work in practice. Typically, the initial step in any international crypto flow involves exchanging fiat currency for a crypto asset. The dominant

[8] Cruz et al. (2024) document a reduction of USDC liquidity in Decentralized Finance protocols in the wake of the crisis.

channel for this is the use of centralized exchanges (CEXs) such as Binance or Coinbase. Users transfer fiat currency to a CEX, which then facilitates the conversion into a crypto asset—for example a stablecoin like Tether's USDT or Circle's USDC—either by acting as the counterparty or by matching the user with another user in a peer-to-peer transaction. These initial transfers of money to the CEX occur within the traditional fiat banking system, and therefore, we do not observe them in our data.

Once a user acquires a stablecoin, they may engage in different activities. For example, they could send the stablecoin to another user on the same CEX. As the stablecoin remains within the same CEX, no on-chain transaction takes place and we do not observe such transactions in our data. Arguably, this is a traditional fiat-based capital flow facilitated by the CEX and not a crypto capital flow. Another possibility is that the user transfers the stablecoin from one CEX to another CEX (and from there, exchanges for fiat and possibly withdraws into a bank account). The transfer of assets between CEXes is recorded on-chain and thus observable. Further, the user might withdraw the stablecoin to a self-custodial wallet for use in another region, which is also observable; or the user could withdraw the stablecoin to a self-custodial wallet and subsequently send it to another self-custodial wallet or CEX, which is recorded on-chain and thus observable.

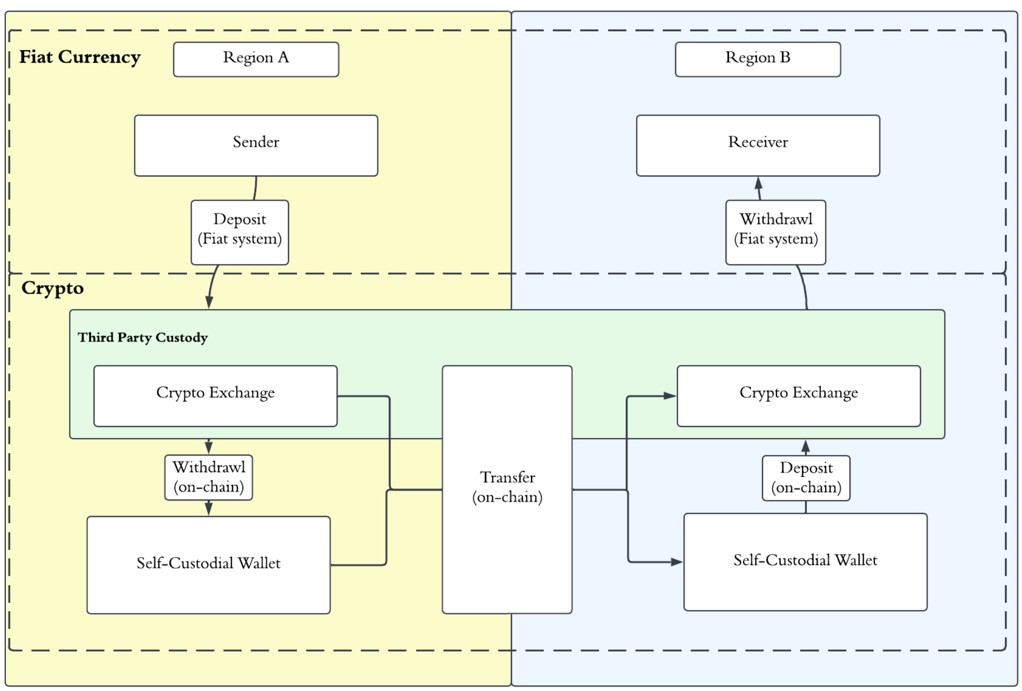



Figure 1: Schematic Overview of International Crypto Flows

Finally, the CEX that has received the fiat money and exchanged it for a stablecoin, has to decide what to do with the fiat money. They could decide to hold on to the fiat money, or to exchange it for another currency (e.g., back into dollars) and tranfer it internationally through the traditional financial system. None of this is observable in our data.

# 3 Data

**Blockchain data.** We obtained full copies of the Ethereum, Binance Smart Chain, Optimism, Arbitrum, Base and Linea blockchains by synchronizing a node with the respective network and extracted the data using the Ethereum ETL package for Python (Medvedev and the D5 team (2018)). For every blockchain, the data spans the period from the genesis block[9] until the last block that has been recorded in 2024. This provides us with data spanning from July 30th 2015, when Ethereum began operating, until December 31st 2024, with other blockchains joining throughout the sample period. The data is transaction level data, exceeding 12 billion transactions and more than 20 terabytes of storage. A transaction occurs when a wallet interacts with the blockchain, such as when transferring crypto assets or interacting with smart contracts.[10] This data also includes all transfers of stablecoins, which we use to map stablecoin flows after estimating the geographic region of wallets.

Further, we combine the transaction level data from the different blockchains in our sample using the fact that wallet addresses carry over between different EVM-compatible[11] blockchains. That is, if we observe that the same wallet address executed transactions on different blockchains, we can be certain that those have been initiated by the same user. This allows us to track the same wallet throughout the whole data set, no matter the specific blockchain a transaction took place on.

**Domain name data.** A wallet address–for the blockchains in our sample–is a 42-character hexadecimal string that starts with '0x', such as `'0xd8dA6BF26964aF9D7eEd9e03E53415D37aA96045'`. Because these strings are cumbersome to handle, domain name systems have been developed that allow users to effectively replace these strings with human-readable domain names, such as 'vitalik.eth'.[12] Users can purchase these domain names from service providers, with the most popular being the *Ethereum Name Service (ENS)*. Purchase prices for these domains typically range from single to triple digit dollar amounts. These purchases are recorded as on-chain transactions and thus data for all purchases is public. For simplic-

[9] The genesis block refers to the first block of a blockchain.

[10] Technically speaking, a transaction is issued by a wallet whenever a user wants to change the state of the blockchain. A transaction can, but does not necessarily have to, include the transfer of assets.

[11] The Ethereum Virtual Machine (EVM) is a virtual machine for executing code that has been adopted by other blockchains to guarantee compatibility.

[12] vitalik.eth is an ENS domain owned by Vitalik Buterin, co-founder of Ethereum.

ity, we obtain our domain name data through queries in Dune Analytics.[13]

**Supplementary data from Dune Analytics.** We also obtain some further supplementary data from Dune Analytics. First is a list of wallets that belong to centralized exchanges such as Binance or Coinbase, totaling 10,072 wallets among 333 entities. Using this data, we identify further wallets belonging to exchanges in a procedure described in appendix C.

Second, we obtain a list of wallets that belong to certain types of bots (i.e. MEV bots). We exclude these wallets because our method would likely mispredict their geographic location, as they are systematically different from the wallets in the training data set that we build. Third, is data on smart contracts for the different blockchains, including the contract addresses and a categorization of the type of contract (e.g., NFTs or decentralized exchanges), and finally, we use data on different stablecoins (e.g., their contract addresses) on different blockchains.

## 3.1 Descriptive Statistics

**Description of blockchain data.** Ethereum is the longest running blockchain in the sample, with its genesis block on July 30th 2015, while the genesis of Linea has been the most recent on July 6th 2023. The latest blocks for the respective chains are both determined by

| **Blockchain** | **Genesis Block** | **Latest Block** | **# Transactions** |
|---|---|---|---|
| Ethereum | 2015, Jul 30 | 21,525,890 | 2,639,611,278 |
| Binance Smart Chain | 2020, Aug 29 | 45,369,482 | 6,523,262,103 |
| Optimism | 2021, Jan 14 | 130,045,411 | 432,212,106 |
| Arbitrum | 2021, May 28 | 290,687,173 | 1,222,934,534 |
| Base | 2023, Jun 15 | 24,450,126 | 1,419,068,646 |
| Linea | 2023, Jul 06 | 14,022,234 | 240,821,235 |
| **Total** | | | **12,477,909,902** |

Table 1: Overview of Blockchain Data

how long the blockchain has been operational and by how fast it produces blocks. Therefore, even though Ethereum is the longest running blockchain in the sample, it has produced less blocks than more recent blockchains such as Optimism or Arbitrum, as it produces them at a slower pace.[14] The number of transactions for each blockchain ranges from hundreds of millions to several billion, with the total number exceeding 12 billion transactions.

[13] Dune Analytics is a commercial blockchain data provider.

[14] For example, Ethereum produces a new block approximately every 12 seconds, while Arbitrum produces a block every 0.25 seconds.

**Description of domain name data.** In total, we obtained almost 6 million domain names, the majority of of which are ENS domains. A detailed breakdown of the domain name data is provided in the following table:

| **Name Service** | **# of Registered Domains** |
|---|---:|
| Ethereum Name Service | 3,413,426 |
| Uxlink | 875,564 |
| Spaceid | 474,021 |
| Basenames | 437,329 |
| Linea ENS Subdomains | 429,956 |
| Arbid | 303,632 |
| **Total** | **5,933,958** |

Table 2: Overview of Domain Name Services

# 4 Methodology for estimating the geographic region of self-custodial wallets

We divide the world into five regions: Africa and the Middle East, Asia and the Pacific, Europe, North America, and Latin America and the Caribbean. To estimate the geographic region of self-custodial wallets, our methodology involves obtaining geographic information for a subset of wallets through two distinct approaches. First, we leverage domain names assigned to wallets purchased through systems such as the Ethereum Name System (ENS). We employ a LLM to infer linguistic and cultural markers—such as language, script, or regional references—that suggest a wallet's likely region. Second, we identify wallets that frequently transact with centralized exchanges targeting specific regional markets, assuming that a wallet predominantly interacting with, for example, a Latin American-focused exchange is likely from that region. These two methods provide an ad hoc regional classification for a subset of wallets, which we then use as labeled training data to train a machine learning model for classification of arbitrary wallets.

Specifically, we train a Gradient Boosted Decision Tree (for background on the method see e.g., Hastie et al. (2017)) leveraging the *Yggdrasil Decision Forests (YDF)* library in Python (Guillame-Bert et al. (2023)) to classify the region of a wallet by recognizing patterns that are characteristic for the respective regions. For this, we construct features capturing wallets' behavioral and transactional characteristics, including time-of-day activity patterns, (non)adherence to daylight savings time, interactions with certain centralized exchanges, and engagement with popular ERC-20 tokens and smart contracts. By learning region-specific patterns the trained model can estimate the geographic region of any arbitrary self-custodial wallet.

## 4.1 Generating training data

### 4.1.1 By analyzing domain names

A significant advantage of using a large language model for inferring the region, country, or language of Ethereum Name System domains is its ability to scale, as it can analyze millions of domains quickly and efficiently. Moreover, ENS supports Unicode Technical Standard 51, which allows users to register domains that employ a diverse range of characters—including those from Arabic, Chinese, Korean, Japanese scripts, and even emojis. This facilitates classification of domain names by providing clear linguistic markers. In some cases, inferring the region can be more straightforward than pinpointing a specific country, as demonstrated by instances where regional characteristics are more pronounced than national ones. However, while making inferences about a country's identity can work particularly well for nations with distinct scripts and large populations—such as China—it tends to be less reliable for smaller countries and countries where linguistic markers may not be as clear-cut.

Despite these advantages, several limitations warrant consideration. One notable challenge is that users might reside in one region while choosing a name that is culturally or linguistically associated with another. For instance, users around the world may select English names or incorporate references to U.S. culture, even if they are not based in the United States. Similarly, migrant populations in Europe or the U.S. might opt for domain names that evoke their country of origin rather than their current locale. Additionally, LLMs can sometimes misinterpret inputs or exhibit biases—for example, when processing German domain names, it may default to associating them with Germany rather than recognizing that they might equally belong to Austria or Switzerland.

For every domain name in our sample, we query the LLM and instruct it to analyze the domain name and guess the country, language, and region the user is from and to provide a short reason for the classification. In essence, our prompt reads:[15]

> "You are trained to classify ENS domain names by country, language, region, and provide a short reason for the classification. Consider references to culture, language, localities, memes. Be creative. Be mindful of the language commonly used in crypto and web3; and of languages that are spoken in many parts of the

[15] The prompt includes some further instructions for the LLM to restrict its answers for the region to the set {Africa and the Middle East, Asia and Pacific, Europe, Latin America and Caribbean, North America, unclassified}.

world, such as English, French and Spanish. If you cannot classify any of these, output 'unclassified' for that particular attribute. Classify the domain: {domain name}."

We restrict the maximum number of tokens of the reply to 150 and set the temperature and "top_p" to 0.4 each. A reply of 150 tokens equals roughly 110-120 words in the English language, while setting the LLM's temperature and top_p to 0.4 makes its responses more deterministic and focused by reducing randomness (temperature) and narrowing the range of token probabilities considered (top_p), leading to safer and more predictable outputs. Some wallets have more than one domain name registered to them. In these cases, we assign the region as the region that was assigned by the majority of the LLM's guesses, excluding "unclassified" responses. That is, a wallet which owns 10 domains, 6 of which have been left unclassified, 3 assigned to Europe and 1 assigned to North America will be assigned to Europe.[16]

#### 4.1.2 By linking wallets to regional exchanges

Some centralized exchanges focus on particular markets that often are restricted to a region or a country. For example, the exchange "Indodax" markets itself as an "Indonesian Bitcoin and Crypto Exchange" and its website is written entirely in Indonesian. We exploit this fact to classify wallets as belonging to a particular region, if they interact particularly frequently with exchanges that focus on that region. The idea is that a self-custodial wallet that frequently receives or sends money to an exchange that is focused on a regional market is likely to belong to a user from that region. Table 12 in the appendix provides a comprehensive list of all exchanges that we classify as particularly regional. We classify a self-custodial wallet as belonging to a particular region, if more than 90% of its transactions with centralized exchanges are with centralized exchanges of that particular region.

### 4.2 Training the classification model

To achieve our goal of estimating the geographic region for any self-custodial wallet, we take the regional classification through domain names and regionally focused centralized exchanges at face value and use them as training data to develop a model that recognizes patterns that are predictive of what region a self-custodial wallet is from. To ensure that we have enough observations on a per wallet basis to reliably recognize patterns, we restrict

[16]Crypto wallets can own multiple domains, similar to how different domain names can lead to the same website. For example, `www.coca-cola.com`, `www.cocacola.com` and `www.coke.com` all lead to the same website.

the sample to only include wallets that have initiated at least 50 transactions. We use the following features to train the classification model:

*Time of Day Features.* For every wallet, we calculate the percentage of transactions that are conducted within every hour of the day. To condense the data and avoid potential overfitting issues, we then estimate a third degree polynomial describing the activity profile and use the coefficients as features.[17]

*Daylight Savings Time Features.* For every wallet, we calculate its activity profile during the months of daylight savings time and during the remainder of the year. This feature exploits regional heterogeneity in the use of daylight savings time. While DST is common in Europe and North America, it is not common in many other regions and non-existent in Asia.

*Top 5 Centralized Exchanges.* For every wallet, we count the number of transactions with a centralized exchange (i.e., withdrawals and deposits). We then sort descending by exchange name and create a categorical variable for the most used centralized exchange, second most used centralized exchange, and so on.

*Top 10 ERC-20 Tokens and Smart Contract Namespaces.* These are two similar features. For the top 10 ERC-20 tokens, we count, for every wallet, the number of transactions using ERC-20 contracts, sort them descending by count and create a categorical variables for the top 10 most used ERC-20 contracts. For smart contracts in general, we count, for every wallet, the number of transactions with smart contracts of the same "namespace"[18], sort them in descending order by count and create categorical variables for the top 10 most interacted namespaces. While the first feature offers some more granularity exploiting heterogeneity on the ERC-20 level, the second feature groups smart contracts based on functionalities and exploits heterogeneity along this dimension.

**Training and Evaluation.** To train and evaluate the model, we split the data into a training data set (90% of the data) and a testing data set (10% of the data). We then train 3 models that we combine to estimate the likely region from the set {Africa and Middle East, Asia and Pacific, Europe, Latin America and Caribbean, North America} that the wallet belongs to. The process is illustrated by the following decision tree:

[17] We use a third degree polynomial to be able to accompany both a minimum and a maximum in the activity profile

[18] The "namespace" groups smart contracts that offer similar functionalities as provided by Dune Analytics.

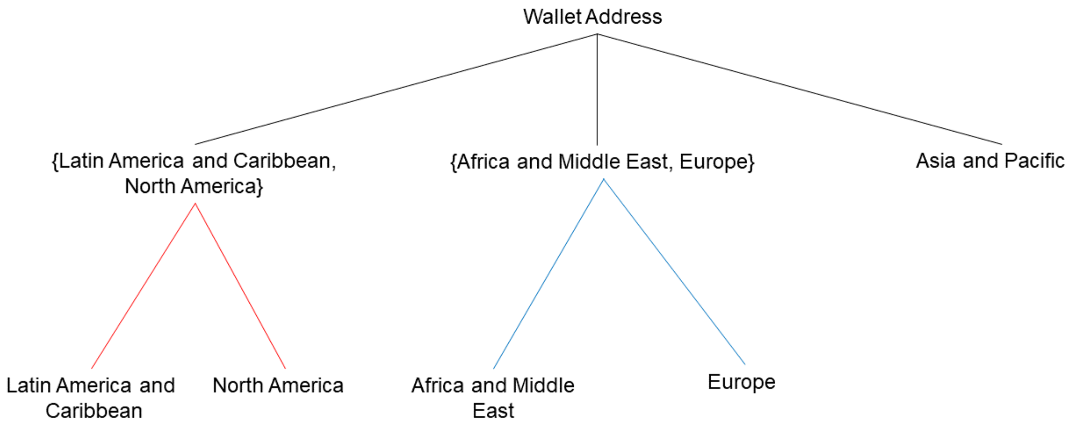



Figure 2: Decision Tree Structure of Classification Models

That is, we first train a model that assigns one of the following three classifications to an address: {North America, Latin America and Carribbean} or {Africa and Middle East, Europe} or Asia and Pacific. Instead of immediately assigning North America, or Latin America and Caribbean, the model assigns the classification that the wallet could be in either region. This approach maximizes the usefulness of the time of day based features, as the grouped regions largely share the same time zones. We then proceed to train two more models. One that splits {North America, Latin America and Carribbean} into its respective regions and one that splits {Africa and Middle East, Europe} into its respective regions. These further models then largely exploit the other features to distinguish between the respective regions. For each of the three models trained, the *YDF* package automatically splits off some of the data for validation to avoid over fitting any given model. To address imbalance in the training data, we calculate balanced sample weights–that is, each observation is weighted inversely proportional to its class frequency–ensuring that underrepresented classes contribute equally to the estimation process (see e.g., King and Zeng (2001), He and Garcia (2009)).

**Classification of Arbitrary Self-Custodial Wallets.** To classify arbitrary self-custodial wallets, we calculate the same features that are used in training the model for wallets that have not been part of the training data and restrict the sample to only include wallets that have at least 50 transactions, to align with the restriction we made in training the model. We then apply the trained model to predict the most likely region the wallet belongs to.

Our *key identifying assumption* is that, conditional on the features used to train the model, wallets of a particular region in the subset that was used to train the model follow the same data generating process as the wallets out of sample that the model is used on to predict regions. For example, we assume that there is no systematic difference between a North American wallet in the subset used to train the model and an out of sample North American wallet with respect to the hours of the day during which they are typically active.

# 5 Results

## 5.1 Training Data

### 5.1.1 Domain Names

Before providing summary statistics of the results, we provide some example classifications of domain names by the LLM to allow the reader to better evaluate the validity of the approach. In Table 3, we present one example for each of the possible regions. The possible region outputs are "unclassified", when the LLM does not find sufficient evidence to assign the domain to a particular region; and "Africa and Middle East", "Asia and Pacific", "Europe", "Latin America and Caribbean", and "North America".

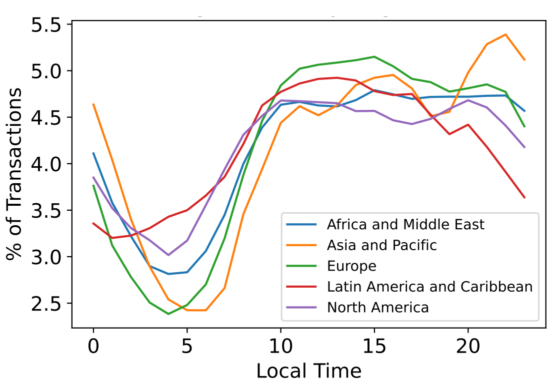
Of course, there is no guarantee that any given guess that the LLM makes is in fact true. Thus, before we use the output of the LLM as input for the training data, we proceed to validate whether this methodology produces reasonable training data. For validation, we collect the timestamps of every single transaction of the wallets that the LLM has classified from our blockchain data. Then, for each wallet, we calculate an activity profile by counting the number of transactions executed within each hour of the day and normalizing by the total number of transactions to derive the % of transactions that have occurred within a particular hour. We then aggregate these wallet-specific activity profiles by region. Since all timestamps in the data are recorded in UTC time, we offset the timestamps for the different regions to convert them to regional timestamps.

For Africa and the Middle East, we use use UTC+2, corresponding to the Central African time zone; for Asia and Pacific, we use UTC+8, corresponding to the time zone that is centered on China, parts of Indonesia, Malaysia, Singapore and Western Australia; for Europe we use UTC+2, which corresponds to Central European Summer Time; for Latin America and Caribbean we use UTC-4, centering on parts of Brazil, Chile and Venezuela; and for North America, we use UTC-6, which corresponds to Central Time in the US. Our intention behind choosing these time zones is not to take a stance on which time zone is the most appropriate to capture an entire region that spans multiple different time zones, but to off-

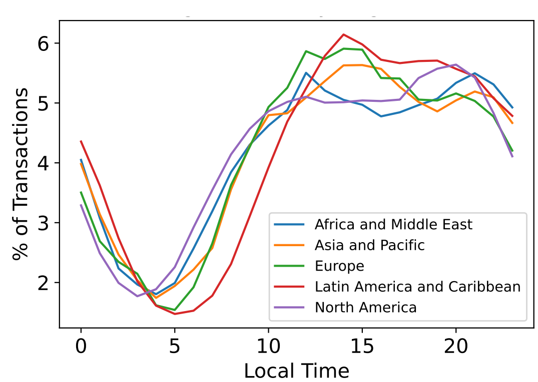
| Name | Region | Country | Language |
|---|---|---|---|
| vucoworld | unclassified | unclassified | unclassified |
| **Reason:** The term ‘vucoworld’ does not clearly reference a specific country, language, or region. It appears to be a coined term possibly related to virtual or digital worlds, which are universal concepts in the crypto and web3 space. | | | |
| الكيماويات | Africa and Middle East | unclassified | Arabic |
| **Reason:** The domain name is in Arabic, which is widely spoken in the Middle East and North Africa, and it translates to ‘chemicals’, indicating a potential relevance to industries in those regions. | | | |
| pijiu | Asia and Pacific | China | Chinese |
| **Reason:** ‘Pijiu’ means ‘beer’ in Chinese, reflecting a popular beverage in Chinese culture. | | | |
| philippzinner | Europe | Germany | German |
| **Reason:** The name ‘philippzinner’ suggests a German origin, likely a personal name, which is common in Germany and associated with the German language. | | | |
| laplazart | Latin America and Caribbean | Mexico | Spanish |
| **Reason:** The name ‘laplazart’ suggests a connection to ‘La Plaza’, which is Spanish for ’The Square’, a common term in many Latin American countries, particularly in Mexico, where plazas are central to community life and culture. | | | |
| lakings | North America | United States | English |
| **Reason:** The term ‘lakings’ likely refers to the Los Angeles Kings, a professional ice hockey team based in Los Angeles, California, which is a significant part of American sports culture. | | | |

Table 3: Example of domain name classifications.

set the UTC based timestamps by a reasonable amount as such to convert UTC time to a reasonable proxy of “local time” for the different regions, balancing out parts of the region that are further ahead or further behind the specific choice of time zone. We then plot the activity profiles that are derived from the LLM’s classification of the wallets as depicted in Figure 3a.




(a) Activity Profile for Wallets Identified by Domain Names



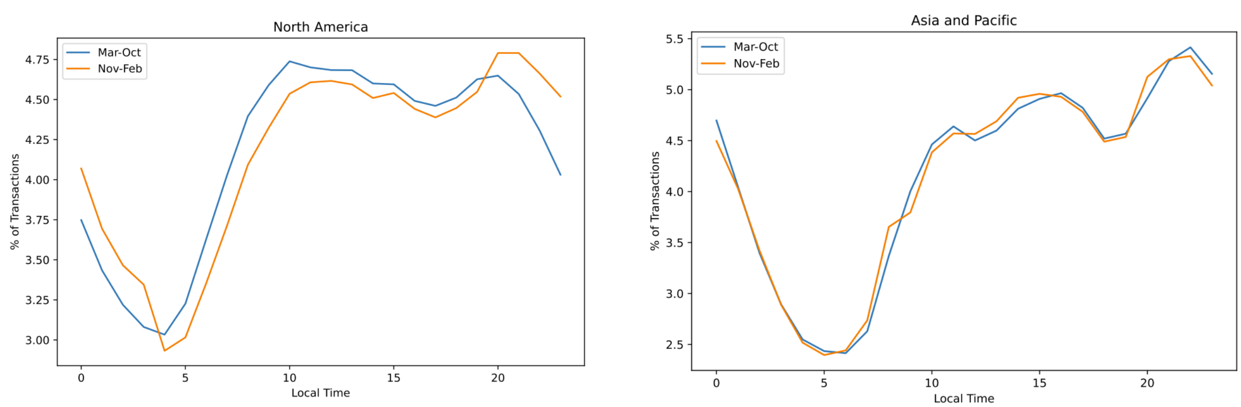

(b) Activity Profile for Wallets Identified through Regional Exchanges

Figure 3: Activity Profiles for Wallets in the Training Data

The activity pattern of all regions in the graph exhibits a distinctive dip in activity during the nighttime, which is intuitive, as we would expect activity to be lower when most people are sleeping. Further, the times of the highest activity span from roughly 10 AM to 10 PM, which correspond to times of the day where most people would be awake. Our interpretation of the figure is that it is supporting evidence that the classification of domain names by the LLM, on aggregate, is successful in assigning the correct regions based on the domain names.

### 5.1.2 Regional Exchange Users

We validate our methodology for identifying wallets that interact with regionally focused CEXs using the same activity profile idea that has been outlined in the section on domain names above. The corresponding activity profiles are depicted in Figure 3b. To offer some further validation of our methodology in creating the training data (both through domain names and regional exchange users), we provide Figure 4 below, that shows that it is also possible to detect the effect of daylight savings time in the classified wallets.




(a) North America

(b) Asia and Pacific

Figure 4: Comparison of the Impact of Daylight Savings Time Across Regions

In the left hand side figure, we plot the activity patterns of wallets that we have identified to be from North America. We fix the x-axis to correspond to the typically assigned time zone for the region (e.g. UTC-6 for North America). Then, we split the data into the time period from March to October, which corresponds to Standard Time in the United States, and the time from November to February, which corresponds to Daylight Saving Time in the United States. The figure clearly shows, that the activity profile in the United States is offset by roughly one hour, as one would expect due to the one hour shift due to Daylight Saving Time. In contrast, the figure on the right hand side shows that there is no significant difference in activity by time of day for the wallets that we have classified to be from Asia and Pacific. This is to be expected, as, with the exception of Australia and New Zealand, no countries in Asia and Pacific observe daylight savings time. We see this as further strong supporting evidence that the classification of wallets in our training data is, on average, correct. Further figures for the other regions can be found in appendix B.

Finally, we present summary statistics of the training data that is generated by our methodology in Table 4. In total, our methodology succeeds in identifying the region of 346,201 wallets, with three quarters being identified through domain names and one quarter through usage of regional CEXs. For domain names, we are able to identify 260,655 wallets out of almost 6 million, that is, roughly 4.4%. While some wallets do carry geographic information in their names, most do not. However, in absolute terms, we succeed in identifying a sizable number of wallets for our training dataset. About half the identified wallets originate from Asia, with the smallest number of identified wallets originating from Latin America and Caribbean. That said, we do not believe that these proportions are necessarily representative of crypto wallets as a whole. Instead, some regions might be easier to identify, for example, due to distinct languages. To account for this imbalance in training the classification model, we employ balanced sample weights as discussed in section 4.2.

| **Region** | **Domain** | **Regional CEX** | **Total** | **% of Total** |
|---|---:|---:|---:|---:|
| Africa and Middle East | 19,159 | 10,466 | 29,625 | 8.6% |
| Asia and Pacific | 124,727 | 49,988 | 174,715 | 50.5% |
| Europe | 90,509 | 2,291 | 92,800 | 26.8% |
| Latin America and Caribbean | 3,723 | 2,090 | 5,813 | 1.7% |
| North America | 22,537 | 20,711 | 43,248 | 12.5% |
| **Total** | **260,655** | **85,546** | **346,201** | **100.0%** |

Table 4: Overview of Training Data

### 5.1.3 Classification Model

After using the training data to train the gradient boosted decision tree classification model, we evaluate the model on the secluded testing data. To describe its performance, we first discuss the confusion matrix in Figure 5 below. It displays the true region of the testing data in the rows and the predicted region of the model in the columns. The data in the matrix has been normalized, such that all rows add up to 100%.

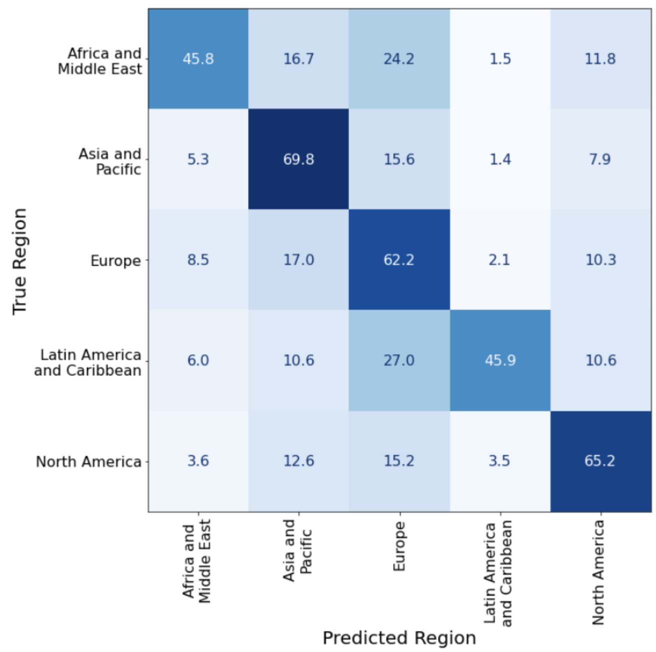



Figure 5: Confusion Matrix of Classification Model

To help interpret the matrix, it is useful to think of some special cases. A model that perfectly classifies the regions would display 100% along its main diagonal, and all off-diagonal entries would be 0. A model that uniformly guesses at random would display 20% in all cells of the confusion matrix. Keeping this in mind, we can see that the model has significant predictive power, performing best in predicting Asia and Pacific (69.8%) and worst at predicting Africa and the Middle East (45.8%). While clearly not perfect, even the predictions for the worst performing region are still significantly better than random guesses. A good performance in prediction indicates two things. First, users in the respective region have relatively distinctive on-chain behavior that the classification model can recognize and use for prediction. Second, it is indicative about the performance of the LLM in classifying the underlying domain names correctly.

To dig further into the model performance, we discuss some of the errors in prediction regions, i.e., the off-diagonal elements of the matrix. If the model makes random errors off-diagonal elements should be symmetric. That is, the proportion of domains from region X that were incorrectly classified as region Y should be equal to the proportion of misclassified domains from region Y as region X. Examining the off-diagonal entries of the confusion matrix, we can see that errors do largely appear to not be random. First, there is a tendency for wallets to be more likely to be misclassified as being in Asia and Pacific, and in Europe. This likely stems from the fact that those two regions contain the majority of observations in the training data, and is a type of bias that we try to avoid by using balanced sample weights. Second, there is a tendency of the model to "misclassify" domains in a fashion that potentially undoes misclassification of the region by the LLM when generating the training data (hence the quotation marks around "misclassify"). To explain this intuitively, note that for example, the proportion of wallets with the true region Africa and Middle East that is classified as North America (11.8%) far exceeds the proportion of wallets with the true region North America that is classified as Africa and Middle East (3.6%). We think that this likely reflects the fact that there is a sizable immigrant population from Africa and the Middle East in North America (who may have chosen domain names that reflect their heritage), which is correctly recognized by the classification model. In contrast, there is not a significant immigrant population from North America that resides in Africa and the Middle East. Therefore, these "misclassifications" of the model are a desirable bias that potentially undoes some incorrect classifications by the LLM when generating the training data. A similar observation can be made for misclassifications of Latin America and North America; and for misclassifications of Africa and Latin America vis-a-vis Europe. With respect to misclassifications of Asia and Pacific, a region that also likely has a significant number of emigration to Europe and North America, we observe that errors are much more symmetric. This is likely due to the fact that the misclassifications towards Asia and Pacific, which stem from it being a majority class, are somewhat balanced by the countervailing bias due to immigration flows.

In appendix D, we provide further details (ROC, AUC, a variety of graphs) about the performance of the three individual classification models that are used.

## 5.2 Estimating the Region of Arbitrary Self-Custodial Wallets

To estimate stablecoin flows, we utilize the predicted probabilities for each region generated by the classification model, rather than relying solely on the most likely predicted region. Incorporating the full distribution of predictions enhances the accuracy and stability of the results. To provide an overview of the total number of self-custodial wallets for which we provide a geographic prediction, we aggregate the regional probabilities. The results are given in Table 5. In total, we have identified 20 million self-custodial wallets that transfer stablecoins on the blockchains included in our sample. We estimate that most wallets belong to Asia and Pacific, with the second most being located in North America. The other three regions have fairly comparable numbers of users. Finally, we manually assigned a handful of wallets that belong to notable entities, as outlined in appendix E.

| Region | # of Wallets | % of Total |
|---|---|---|
| Africa and Middle East | 3,010,170 | 13.59 |
| Asia and Pacific | 6,945,933 | 31.36 |
| Europe | 3,231,460 | 14.59 |
| Latin America and Caribbean | 2,556,156 | 11.54 |
| North America | 4,431,223 | 20.00 |
| **Total** | 20,174,942 | 100.00 |

Table 5: Predicted Regions for Self-Custodial Wallets

# 6 Estimating International Stablecoin Flows

While our data covers the timeframe from the inception of stablecoins up until the end of 2024, we focus on presenting the data for 2024 in this section. First, we present some summary statistics for stablecoin flows in 2024, before showing detailed breakdowns of stablecoin flows in the following subsections. The total number of stablecoin transactions that we map is around 138 million, totalling a volume of $2,019 billion, implying an average transaction size of $14,630.[19] The total number of wallets involved in these transactions is 14.6 million, which can be subdivided into 10.4 million that belong to centralized exchanges and 4.2 million self-custodial wallets. The total volume of $2,019 billion can be split into three categories: $309bn in flows entirely between self-custodial wallets, $1,141bn in flows between self-custodial wallets and centralized exchanges and $569bn in flows between centralized exchanges. This breakdown highlights the importance of determining the geographic region of self-custodial wallets, as they are present in at least one side of the transaction in 72% stablecoin transactions as measured by volume. There is considerable heterogeneity in the regional useage of the different stablecoins, as outlined in Table 6 below.

As shown in the table, USDT is more popular in regions that feature more emerging markets (i.e., Africa and Middle East, Asia and Pacific, Latin America and Caribbean), while USDC is more popular in regions that feature more advances economies (Europe, North America). There is also considerable heterogeneity in the average size of transactions. Intuitively, transaction sizes in regions that are more economically developed also tend to be larger (cf. Table 7) That is, North America has the largest average transaction size at $35,016, with the smallest occurring in Asia and Pacific ($11,493). Further, stablecoin flows exhibit fat tails, with averages significantly exceeding the median.[20]

[19] We exclude stablecoin transactions where the sender and receiver is the *same* centralized exchange, as these are typically due to operational needs. That is, they are flows from deposit addresses to hot wallets and flows between hot wallets and cold storage wallets.

[20] We exclude stablecoin flows with values less than 1 cent in the calculations to avoid skewing them due to so called "dusting attacks".

| Region | Stablecoin | Volume (USD bn) | Percentage |
|---|---|---|---|
| Africa and Middle East | USDC | 85 | 42.7 |
| Africa and Middle East | USDT | 115 | 57.3 |
| Asia and Pacific | USDC | 179 | 42.0 |
| Asia and Pacific | USDT | 247 | 58.0 |
| Europe | USDC | 167 | 50.0 |
| Europe | USDT | 167 | 50.0 |
| Latin America and Caribbean | USDC | 68 | 43.3 |
| Latin America and Caribbean | USDT | 88 | 56.7 |
| North America | USDC | 273 | 61.3 |
| North America | USDT | 172 | 38.7 |

Table 6: Volume and Percentage of Stablecoins by Geographic Region

| Region | Average Transaction Size | Median Transaction Size |
|---|---|---|
| Africa and Middle East | $13,108 | $100 |
| Asia and Pacific | $11,493 | $94 |
| Europe | $18,878 | $200 |
| Latin America and Caribbean | $14,005 | $51 |
| North America | $35,016 | $101 |

Table 7: Average Transaction Sizes by Region

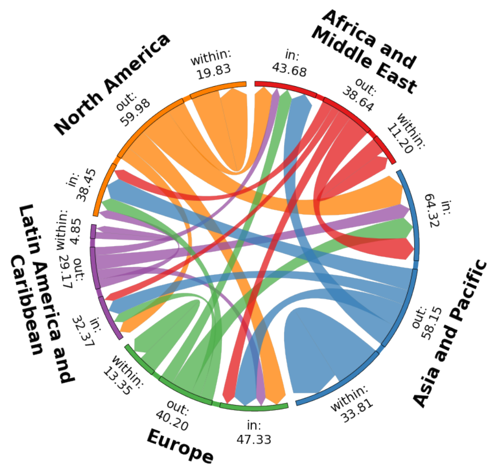
Further, we present regional heterogeneity in the preference for interacting with Binance and Coinbase, as outlined in Table 8. The popularity of Binance and Coinbase across regions follows a similar trend to the popularity of USDT and USDC across regions. That is, regions with more emerging markets tend to favor interactions with Binance, while Coinbase's popularity increases in regions with more advanced economies.

| Region | Coinbase (% Volume) | Binance (% Volume) |
|---|---|---|
| Africa and Middle East | 25.7 | 74.3 |
| Asia and Pacific | 16.9 | 83.1 |
| Europe | 33.7 | 66.3 |
| Latin America and Caribbean | 27.7 | 72.3 |
| North America | 54.0 | 46.0 |

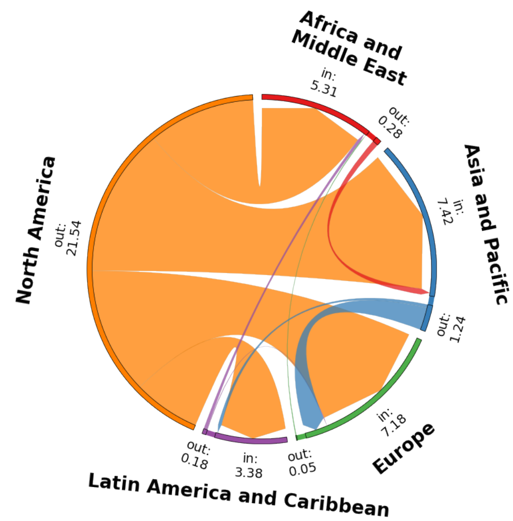
Table 8: Percentage of Volume with Coinbase and Binance by Region

## 6.1 Stablecoin Flows Between Self-Custodial Wallets

This section focuses on the description of stablecoin flows between self-custodial wallets of different geographic regions. In Figure 6, we present our estimates of both gross and (bilateral) net flows of stablecoins in 2024 between regions.




(a) Regional Stablecoin Gross Flows




(b) Regional Stablecoin Net Flows

Figure 6: Regional Stablecoin Gross and Net Flows
(in billion dollars)

We begin with the figure displaying gross flows on the left hand side. For each region, we split flows into three categories. Inflows, outflows and within flows (where both the sender and receiver are in the same region). The largest regions in terms of stablecoin flows are Asia and Pacific ($156.28 bn), followed by North America ($118.26 bn), while Latin America and Caribbean has the smallest stablecoin flows ($66.39 bn). However, relative to GDP, both Africa and Middle East (1.5%) and Latin America and Caribbean (1.4%) exhibit significantly higher stablecoin flows than the other regions (Asia and Pacific: 0.4% , Europe 0.4%, North America: 0.4% ), hinting at the popularity of stablecoins and crypto assets, particularly in emerging markets, that has been described in the literature (see e.g., von Luckner et al. (2023), Cardozo et al. (2024), Cerutti et al. (2024)).[21]

As a general pattern, within region flows are sizably smaller than between region flows. This is in line with the idea that stablecoins are particularly attractive for international payments and remittances, an area in which traditional transfers are particularly slow and costly (see e.g., World Bank (2024)). In contrast, within region (and in particular, within country) payments are typically significantly more efficient than their international counterparts, making stablecoins less attractive as an instrument to conduct such payments. Further, within country payments are also more likely to be denominated in local currency

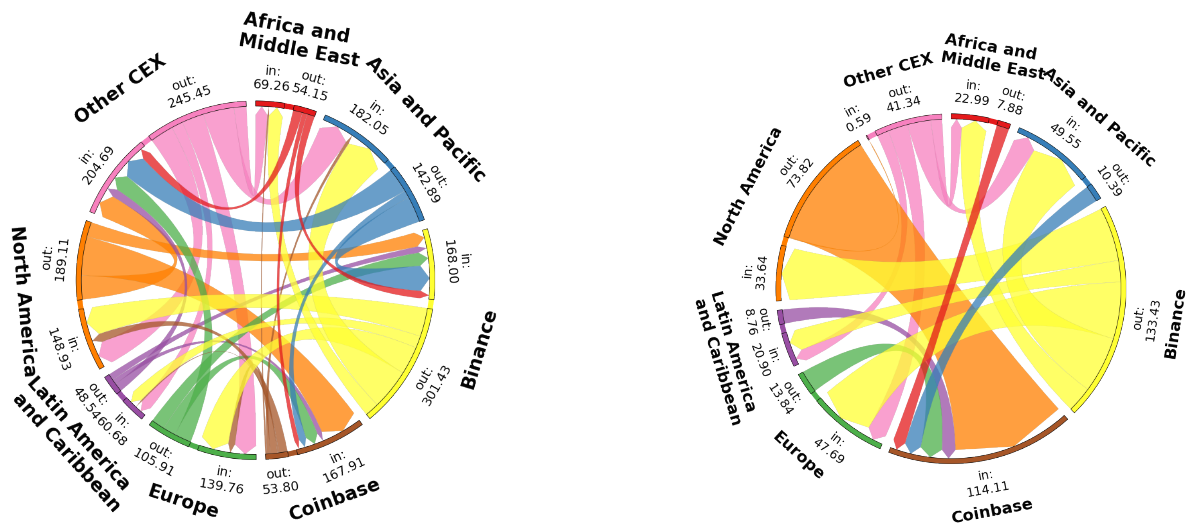
[21]Percentages calculated as (inflows + outflows + within flows)/GDP. Regional GDP numbers aggregated from country level World Economic Outlook 2023 data (in trillion dollars): Africa and Middle East: 6.31, Asia and Pacific: 38.35, Europe: 27.1, Latin America and Caribbean: 4.89, North America: 33.35

rather then dollars, contributing to limited utility of stablecoins for domestic payments in economies that are not dollarized. Finally, on net flows, virtually all stablecoin net flows are outflows from North American wallets ($21.54 bn) to the other regions.

## 6.2 Stablecoin Flows Between Centralized Exchanges and Self-Custodial Wallets

In this section, we describe stablecoin transfers between self-custodial wallets and centralized exchanges. Centralized exchanges take a pivotal role in the crypto ecosystem, as they offer the most prominent venues for users to exchange fiat currencies and crypto assets. That is, when a user wants to buy a crypto asset with fiat currency, typically the transaction is intermediated by a centralized exchange. Similarly, when a user wants to sell a crypto asset for fiat currency, the transaction is typically intermediated by a centralized exchange.

In Figure 7, we present our estimate of stablecoin flows between centralized exchanges and the geographic regions. For centralized exchanges, we separately depict Binance and




(a) Gross Flows

(b) Net Flows

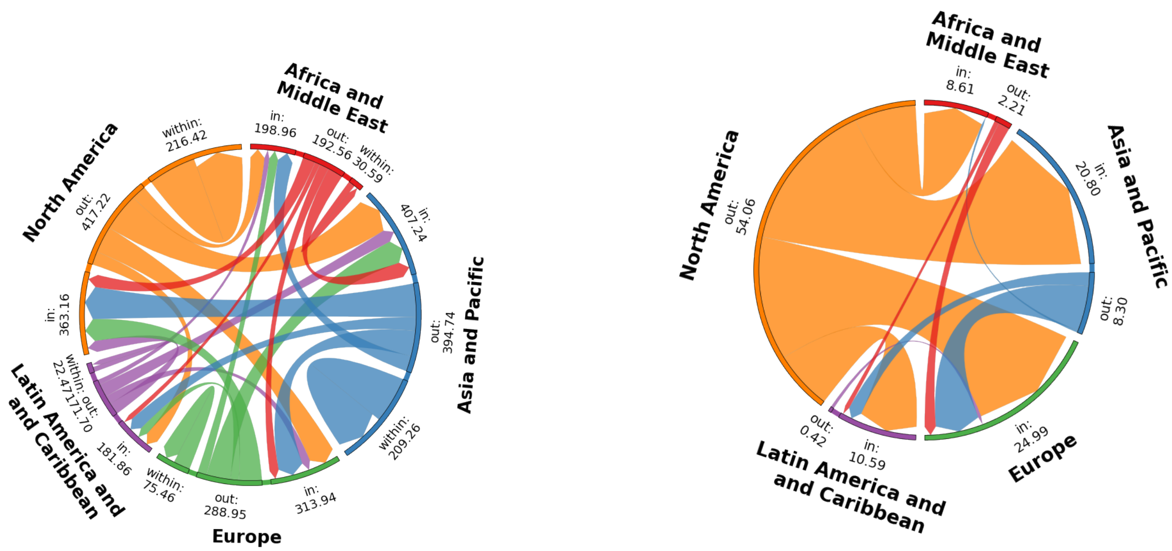
Figure 7: Stablecoin Flows between Self-Custodial Wallets and Centralized Exchanges (in billion dollars)

Coinbase as the two individual exchanges that are involved in the largest amount of flows. All other exchanges are grouped into the "Other CEX" category. In terms of volumes, the critical role of centralized exchanges in bridging the fiat and crypto system becomes apparent. Binance alone processes more volume than any single geographic region of self-custodial wallets. As a general pattern, Binance and the Other CEX facilitate the vast majority of inflows of stablecoins into self-custodial wallets for all geographic regions. In contrast, Coinbase receives far more inflows of stablecoins from self-custodial wallets than outflows. This pattern hints at the possibility that Binance and other exchanges typically facilitate "on-ramping", that is the exchange of fiat currency for stablecoins, some of which are then further transferred into self-custodial wallets. In contrast, Coinbase may be used more frequently by users wanting to "off-ramp", that is, exchanging stablecoins back into fiat currency. This may also be an expression of the geographic focus of different exchanges. While Binance has stronger ties to regions with more emerging economies, where capital flight motives may be more acute, Coinbase has stronger ties to the American market, and might be better suited for transferring capital into the American financial system.

## 6.3 Dividing Centralized Exchange Flows into Regional Flows

In this section, we enhance our analysis of region-to-region stablecoin flows by assigning a regional estimate for flows involving centralized exchanges. To achieve this, we allocate each centralized exchange's stablecoin flows to specific geographic regions based on the proportion of its stablecoin volume associated with each region. Specifically, for each exchange, we calculate the percentage of its total stablecoin volume that is transacted with self-custodial wallets in each region. We then use these percentages to distribute all flows involving the exchange among the regions. This method relies on the assumption that the geographic distribution of self-custodial wallets interacting with the exchange approximates the geographic distribution of the exchange's users. A key advantage of this approach over the web traffic-based assumption employed by Chainalysis is that it does not assume that CEX users do not use VPNs and that the average transaction sizes across regions are uniform. In contrast, by utilizing volume data, our method inherently accounts for regional differences in transaction sizes. The resulting allocation of stablecoin flows is illustrated in Figure 8. The left-hand side figure illustrates gross stablecoin flows, while the right hand-side illustrates net flows. The relative breakdown of the flows is very similar to the breakdown of flows between self-custodial wallets depicted in 6. This indicates that the interactions of self-custodial wallets with CEXs, which we have used to assign the CEX flows, is not meaningfully different from the interactions of self-custodial wallets with self-custodial wallets.

As before, North America and Asia and Pacific are the largest regions in terms of absolute stablecoin flows, while Africa and Middle East, and Latin America and Caribbean exhibit larger flows relative to GDP (6.7% and 7.7% respectively). Inter-region flows again are more important than within-region flows, supporting the narrative that stablecoins are used for international capital flows and remittances.




(a) Regional Stablecoin Gross Flows

(b) Regional Stablecoin Net Flows

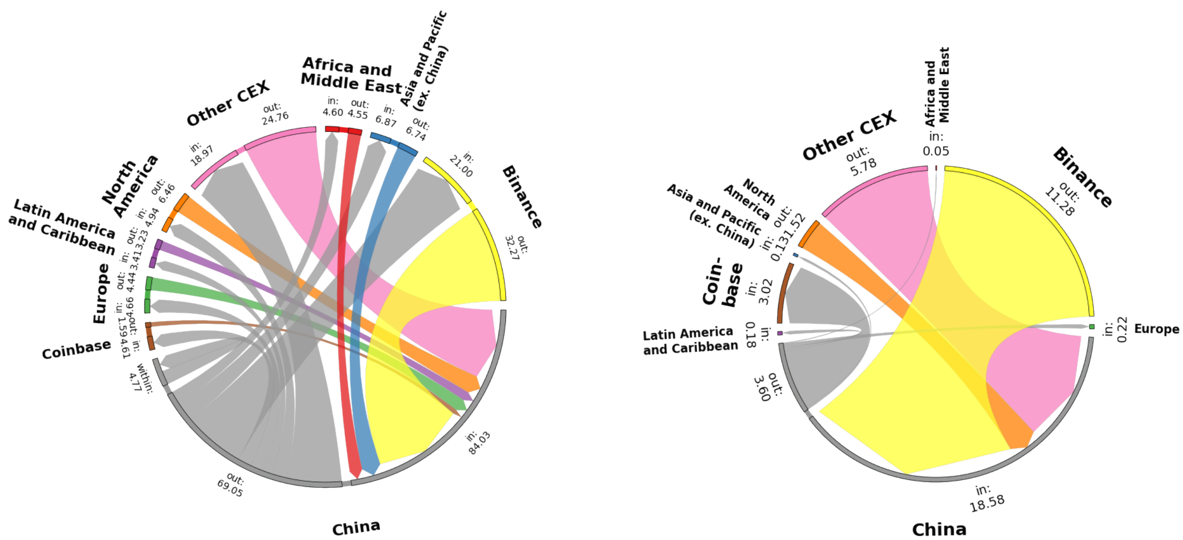
Figure 8: Regional Stablecoin Gross and Net Flows
(in billion dollars)

For net flows, the most significant flow is an outflow from North America totalling \$54.06bn, flowing into all other regions, satisfying some of the global demand for dollars. Furthermore, the operational mechanics of stablecoin issuance may structurally position North America as a natural “exporter” of stablecoins. Specifically, (eligible) users—likely more often located in North America due to greater access to fiat dollars—transfer \$1 in fiat currency to issuers, who then mint an equivalent value in stablecoins. Price stability is maintained through arbitrage: when a stablecoin’s market price exceeds \$1, North American participants can profitably exchange fiat dollars for newly issued stablecoins and sell them on secondary markets, reinforcing the region’s net outflow position (see Makarov and Schoar (2022)).

## 6.4 Country Focus: China

Within this section, we take our methodology one step further to offer country level stablecoin flow analysis for China. The reason we take a closer look at China is two-fold. On the one hand, China has anecdotally been described as very active in crypto assets, even though they have officially been banned. On the other hand, current estimates about international crypto flows involving China primarily stem from use of the Chainalysis dataset, which assumes that, on average, users do not use VPNs when interacting with CEXs. An assumption that is likely systematically violated when it comes to China. For example, the website Binance.com is not accessible from China without use of a VPN. In fact, our analysis suggests that Binance is *the* most important CEX in China in terms of stablecoin flows, despite being accessible by VPN only.

To distinguish self-custodial wallets between China, and Asia and Pacific (ex. China), we train an additional model, which achieves 79% accuracy in separating wallets into these two categories.[22] We present all 2024 stablecoin flows that involve China as the receiver or sender in Figure 9. We begin by describing gross flows on the left hand side of the figure. We




(a) Gross Flows

(b) Net Flows

Figure 9: Stablecoin Flows Involving China (in billion dollars)

estimate that stablecoin flows involving China are sizable, with inflows totaling \$84.03bn, outflows of \$69.05bn, and within country flows of \$4.77bn. Given the small amount of within-flows relative to in- and outflows, stablecoin flows involving China seem to mostly facilitate international capital flows, rather than domestic payments. The most important counterparty for Chinese self-custodial wallets is Binance, facilitating \$32.27bn in flows into Chinese self-custodial wallets, and absorbing \$21.00bn in flows from them. It is notable that the category of "Other CEX" also plays a significant role, while gross flows involving Coinbase are relatively small. In terms of geographical regions, Asia and Pacific (ex. China) is the most important counterparty, with relatively balanced in- and outflows of \$6.87bn and \$6.74bn respectively. In terms of net flows, we estimate bilateral net inflows of \$18.58bn into

[22]We offer more detailed evaluation metrics in appendix D.

Chinese self-custodial wallets, stemming from Binance ($11.28bn), Other CEXs ($5.78bn) and North America ($1.52bn). As a percentage of the current account surplus, we estimate the net inflows of stablecoins into China to amount to 4.4%.[23] Outflows from Chinese self-custodial wallets mostly flow to Coinbase ($3.02bn).

# 7 Economic Drivers of Stablecoin Flows

We further validate our international stablecoin flow estimates by linking them to exchange rates, which are closely related to some economic drivers (e.g., inflation) that have been documented by previous research (e.g., Auer et al. (2025)), and also provide further validation—together with a novel result—by documenting the sensitivity of international stablecoin flows to systemic shocks, such as the March 2023 U.S. banking crisis.

## 7.1 Link between Stablecoins and Exchange Rates

Stablecoins are typically minted in the U.S., where issuers convert fiat dollars into digital tokens. Taken together with our previous analysis that showed that stablecoin net flows are largely outflows from North America. We thus hypothesize that net outflows from North America to other regions could be linked to demand for U.S. dollars, for example through the exchange rate vis-à-vis the US dollar due to a desire to hold US-dollars when local currencies depreciate. We use our dataset to generate a panel time series of daily net stablecoin flows from North America to the other geographic regions from January 1, 2022, to December 31, 2024.[24] Then, we estimate the following regression:

$$\text{Net Flows vs NA}_{r,t} = \beta_1 \text{VIX}_t + \beta_2 \text{Broad Dollar}_t + \beta_3 \text{Crypto F\&G}_t + \alpha_{r,Q} + \text{Weekend} + \epsilon_{r,t}$$

Here, Net Flows vs $\text{NA}_{r,t}$ denotes net flows between North America and region $r$ on day $t$. The VIX index is a common measure of market volatility, while the Broad Dollar Index is a trade-weighted index of the exchange rates between the US-Dollar and several other currencies. The Crypto Fear & Greed Index (Crypto F&G) captures sentiment in crypto markets, with higher values capturing bullish sentiment, and lower values capturing bearish sentiment. We include region-by-quarter fixed effects ($\alpha_{r,Q}$) to account for regional and

[23] We use the current account surplus of $424bn in 2024 from the IMF's World Economic Outlook in this calculation.

[24] We begin the sample period in 2022, as stablecoins had largely established themselves by that point in time. In the time prior, stablecoins had been relatively new and unestablished, likely exhibiting different behavior. For example, the market capitalization of USDT grew from $22bn to $78bn in 2021, while USDC grew from $4bn to $42bn.

temporal variation, and a weekend dummy to control for cyclical patterns—stablecoin flows tend to be significantly lower during weekends. For the VIX and the Broad Dollar Index, weekend values are interpolated linearly to align with the 24/7 nature of stablecoins and to avoid dropping weekend observations.

**Results.** Table 9, left column, shows that a stronger U.S. dollar (higher Broad Dollar Index) is associated with a significant increase in outflows of stablecoins from North America into other regions.[25] In contrast, we do not find a significant impact of volatility or crypto market sentiment on net stablecoin flows from North America to other regions.

| | Global | China |
|---|---|---|
| VIX | -0.018 | 0.098 |
| | (0.068) | (0.053) |
| Broad Dollar | 0.181** | |
| | (0.085) | |
| USD/CNY | | 0.284*** |
| | | (0.095) |
| Crypto F&G | 0.025 | -0.040 |
| | (0.042) | (0.045) |
| Region × Quarter FE | ✓ | |
| Quarter FE | | ✓ |
| Weekend FE | ✓ | ✓ |
| Observations | 4·1096 | 1096 |
| R-squared | 0.310 | 0.200 |
| F-statistic | 40.580 | 16.813 |

Standard errors in parentheses
$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$
Region: Standard errors clustered at the time level.
China: Robust standard errors.

Table 9: Panel Regression for Drivers of Net Stablecoin Flows

Note: Residuals in both regressions have been tested to confirm stationarity at the 0.01 level. A further specification that includes a lag of the flows into the regression can be found in Appendix F.

Next, we extend this analysis to the country level at the example of China. For this, we create a time series using our China specific estimates, as outlined in section 6.4. The data consists of daily observations from January 1, 2022 to December 31, 2024. In addition to net flows between North America and China, we consider net flows between Binance and China, as Figure 9b indicates that Binance is the most important centralized exchange in

[25]While we cannot attribute causality in the relationship between international stablecoin flows and exchanges rates, reverse causality—i.e., stablecoin flows driving exchange rates—seems unlikely given the relative size of stablecoin flows to overall capital flows.

facilitating net stablecoin flows into Chinese self-custodial wallets. We estimate the following regression:

$$\text{Net Flows vs (NA+Binance)}_t = \beta_1 \text{VIX}_t + \beta_2 \text{USD/CNY}_t + \beta_3 \text{Crypto F\&G}_t + \alpha_m + \text{Weekend} + \epsilon_t$$

**Results.** Table 9, right column, confirms that a stronger dollar versus the Chinese Renminbi (i.e., higher USD/CNY) is associated with increased stablecoin flows into China. As in the region-level regression, we also do not find a significant association with volatility as measured by the VIX, or of crypto market sentiment (e.g., the Crypto Fear&Greed index) for China.

## 7.2 The Disruption of Stablecoin Flows in the March 2023 Banking Crisis

Next, we further validate our dataset and present novel evidence of the impact of the March 2023 US banking crisis on stablecoin flows. Stablecoin issuers are in the business of intermediating between fiat money and stablecoins. For this, they rely on banks to manage fiat reserves and to settle the fiat leg of stablecoin issuance and redemption transactions. During stablecoin issuance, customers have to send fiat to stablecoin issuers, while during redemption stablecoin issuers have to return fiat to their customers—transactions that have to be facilitated by banks. The March 2023 banking crisis exhibited failures of several crypto-related institutions (Silicon Valley Bank, Signature Bank, Silvergate) that provided banking services to both stablecoin issuers and CEXs. When these banks failed, the fiat currency side of stablecoin issuance was disrupted. As a consequence, we hypothesize that this shock caused a distruption in stablecoin flows originating from North America. Using a difference-in-differences approach, we compare weekly stablecoin flows originating from North America (i.e., the treated group) to weekly flows originating from other regions (i.e., the control groups) over five weeks pre- and post-crisis. We set the treatment day to be March 10th 2023, the day Silicon Valley Bank collapsed and was put under FDIC receivership. We estimate the event study style differences in differences regression:

$$\text{Flows from Region}_{r,t} = \alpha_r + \alpha_t + \sum_{\substack{\tau=-5, \\ \tau \neq -1}}^{\tau=5} \beta_\tau \cdot \text{Treated}_r \cdot 1\{\tau = t\} + \epsilon_{r,t}$$

**Results** (Figure 10, Table 15 in appendix F) show parallel pre-trends, highlighting that the other geographic regions are reasonable controls for North America. On impact of the shock, stablecoin flows originating from North American drop dramatically, exhibiting a decrease of almost 10 standard deviations relative to controls. In the following weeks, flows normalize, returning to normal. This underscores the banking sector's critical role in stablecoin markets and the disruption that the March 2023 banking crisis has caused for stablecoins.

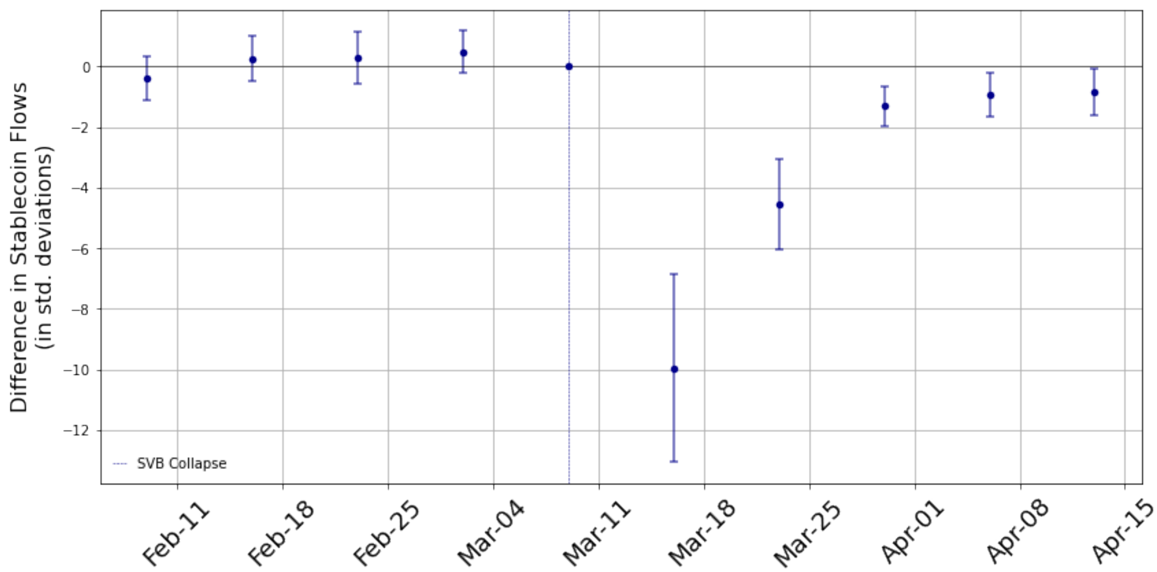



Figure 10: Impact of March 2023 Banking Crisis on Stablecoin Flows

# 8 Comparison with Chainalysis Dataset

In this section we compare the dataset developed in this paper to the commercially available Chainalysis dataset that has been used in several papers (Cerutti et al. (2024), Cardozo et al. (2024), Auer et al. (2025)). Before the comparison, we briefly explain the methodology behind the Chainalysis dataset. To establish the geography of crypto asset flows, Chainalysis focuses on on-chain transactions between CEXs. To infer the geographic location of users, they obtain web traffic data for these exchanges from the commercial provider Similarweb. The idea is that the web traffic data is a good proxy for the geographic location of the exchanges users. Consider the stylized example below, which we borrow from Cerutti et al. (2024):

> "Imagine a stylized example with two exchanges, three countries, and a hypothetical transaction volume of 100 Bitcoin from exchange 1 to exchange 2 on a given day. Based on the web traffic pattern shown in Figure 11, Chainalysis

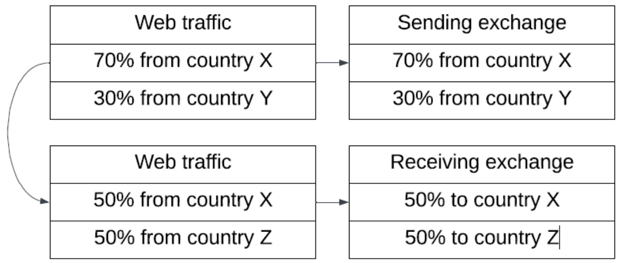



Figure 11: Example from Cerutti et al. (2024).

would distribute this daily transaction volume as follows: 35 Bitcoin from country X to country X, 35 Bitcoin from country X to country Z, 15 Bitcoin from country Y to country X, and 15 Bitcoin from country Y to country Z."

Embedded in Chainalysis methodology are two main assumptions. First, users do not use VPNs, which would possibly falsify their geographic location. Second, users from different countries make, on average, transactions of the same sizes, which is needed to split transaction volumes in equal proportion to web traffic. The dataset derived in our paper relies on neither of those assumptions, which we see as a distinct advantage. In fact, section 6 has provided direct evidence to invalidate the assumption of equal transaction sizing across regions, highlighting that North American stablecoin flows, for example, are on average significantly larger than those of other regions. Further, the data produced by our methodology allows us to map stablecoin flows between region (and sometimes countries) and CEXs (cf. Section 6.2), which is not possible in the Chainalysis dataset. In particular, being able to highlight the crucial role that CEXs play in facilitating international stablecoin flows is a unique advantage of our dataset. However, the method in our paper also has drawbacks. First, there are challenges to classifying wallets using domain names and machine learning, as discussed in section 4. Second, the usage of web traffic data by Chainalysis allows for country level estimates, while our method largely only produces region level estimates.[26] To compare the datasets, we aggregate the Chainalysis dataset from a country level to a region level, by assigning countries to regions as in the IMF's World Economic Outlook. We then compare the two datasets along two dimensions. First, we compare the estimated quantities. Second, we analyze correlations between the two datasets.

## 8.1 Quantities

Prior to conducting a comparative analysis of quantities, it is important to discuss key differences in the coverage of the underlying data. The dataset derived by our methodology

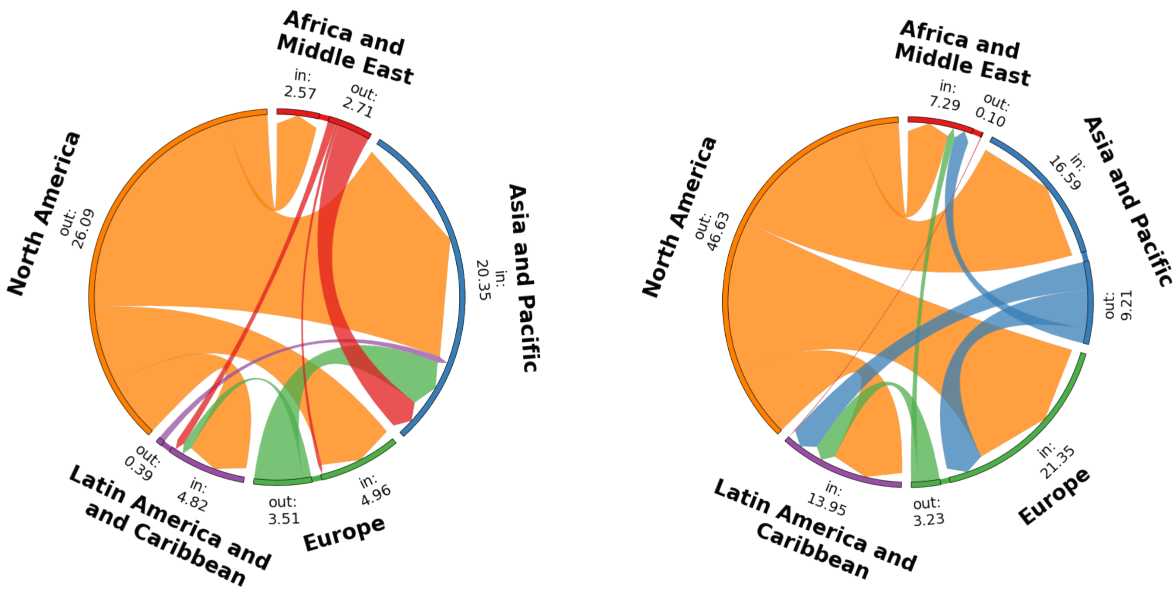
[26]Our methodology can derive country level estimates, such as those for China, on a case-by-case basis.

encompasses a subset of blockchains, whereas the data utilized by Chainalysis is likely more comprehensive in its coverage. However, within the blockchains covered, our dataset is likely more exhaustive. Chainalysis primarily captures transactions that both originate and terminate at CEXs. To extend their coverage, their "indirect" category attempts to trace transactions that pass through self-custodial wallets en route between CEXs. This is likely to only capture a fraction of the transactions involving self-custodial wallets and introduces additional complexities, complicating measurement.

That said, the overall quantities are broadly comparable, with our dataset estimating a total of $2 trillion, while Chainalysis reports a figure of $1.7 trillion. Both datasets agree that USDT exhibits greater prevalence in emerging market (EM) regions, whereas USDC is more favored in advanced economy (AE) regions. Further, there is agreement that Asia and Pacific exhibits the largest stablecoin flows, while Africa and the Middle East, as well as Latin America and the Caribbean, record smaller absolute volumes. Additionally the *direct* category of Chainalysis flows also estimates that net stablecoin outflows predominantly originate from North America, flowing toward all other regions.

The most notable differences relate to the "indirect" category in the Chainalysis dataset, and the estimates for China. To be able to compare the datasets using a direct and and indirect category as is provided in Chainalysis, we construct an analog direct and indirect category for our dataset. For the direct category, we include all direct flows between CEXs, while the indirect category includes all flows that involve self-custodial wallets (i.e., flows between CEXs and self-custodial wallets, and flows between self-custodial wallets). We find no significant disagreement between the direct category in both datasets (cf Figure 12), regarding the observation that net stablecoin flows are largely outflows from North America to other regions.

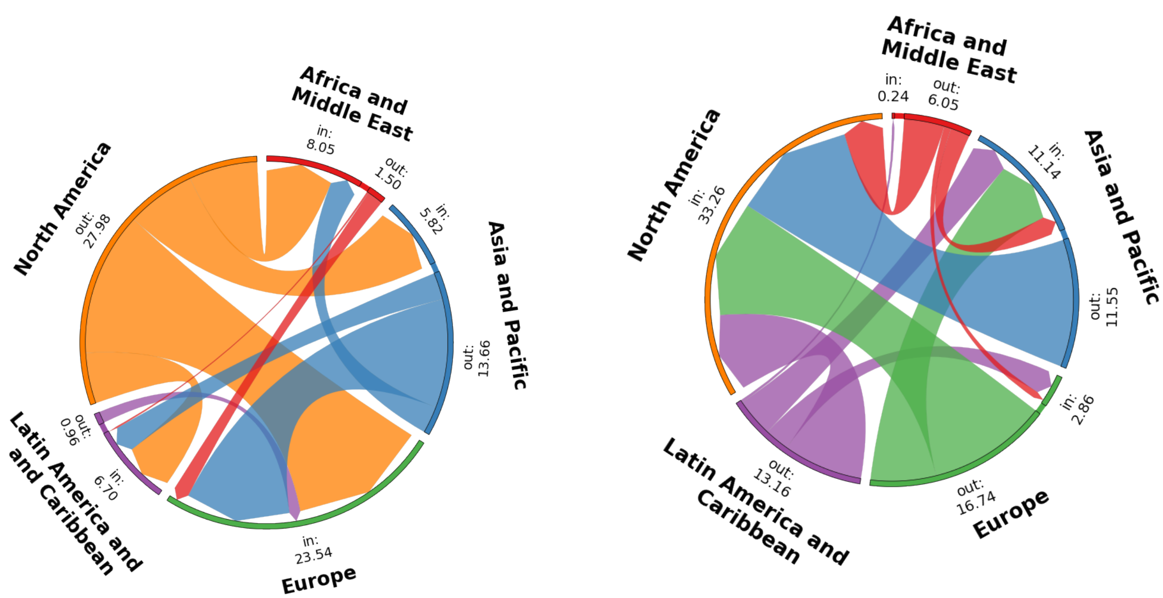
However, there is a severe disagreement between the indirect category in the Chainalysis dataset and our indirect category. (cf. Figure 13). While our indirect category estimates that stablecoins primarily flow from North America to other regions worldwide—just as in the direct category—the indirect category in the Chainalysis dataset suggests the opposite pattern, with stablecoins flowing out of all regions into North America. The disagreement of the indirect category in Chainalysis with our indirect category is puzzling, as both indirect categories likely cover the same underlying transfers of stablecoins between self-custodial wallets and CEXs. It is difficult to think of an obvious explanation that accounts for this discrepancy. Looking at the data as a whole, it is notable that all categories, except for the indirect category in Chainalysis, agree that net flows are largely outflows from North America to other regions. Therefore, the disagreement of the indirect category in Chainalysis with all the other estimates raises questions about its reliability.




(a) "Direct" Net Flows in our Dataset

(b) "Direct" Net Flows in Chainalysis

Figure 12: Comparison of Net Flows in Direct and Indirect Categories Calculated for Our Dataset

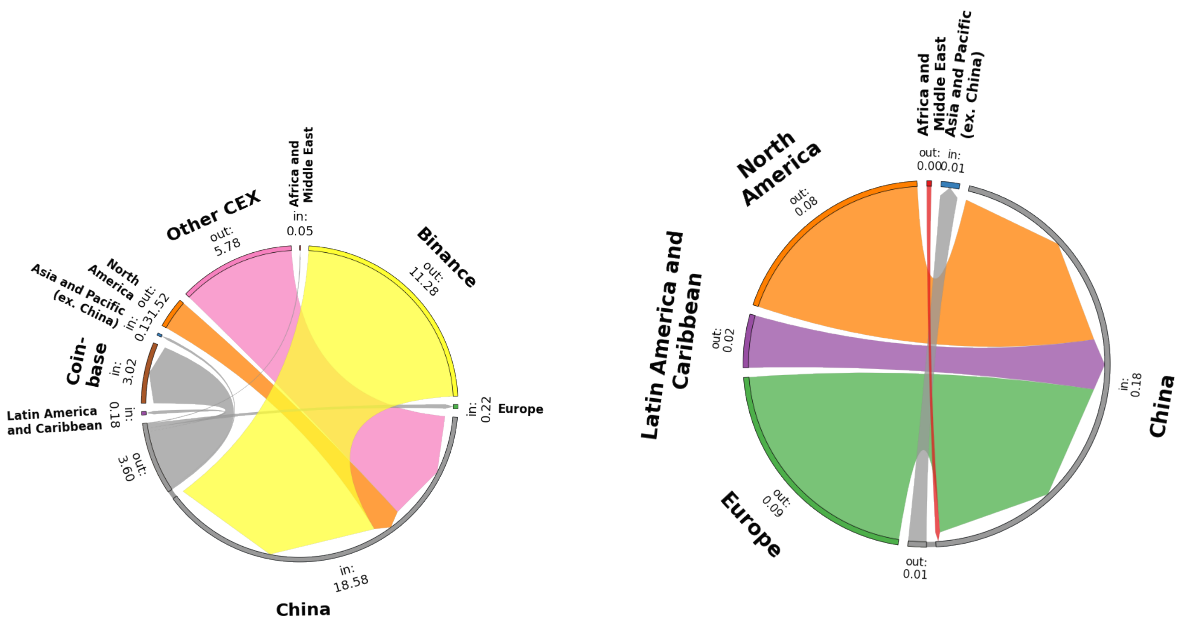



(a) "Indirect" Net Flows in our Dataset

(b) "Indirect" Net Flows in Chainalysis

Figure 13: Comparison of "Indirect" Net Flows between our Dataset and the Chainalysis data

The next significant disagreement between the two datasets is with respect to China. In the case of China, significant use of VPNs likely results in Chainalysis methodology capturing only a fraction of actual activity, which we believe to be reflected in the difference between the estimates. Our estimates indicate gross flows involving China amounting to \$153 billion, 5.5 times larger than the \$28 billion reported by Chainalysis, and net flows of \$18 billion, a staggering 100 times larger than Chainalysis' estimate of \$0.18 billion (cf. Figure 14).[27]




(a) China net flows in our data

(b) China net flows in Chainalysis (Direct + Indirect)

Figure 14: Comparison of net stablecoin flows involving China between both datasets.

## 8.2 Correlations

While the previous section has compared the quantity estimates between both dataset, this section compares the dynamics by analyzing the correlations between the datasets. To highlight connections to the quantity estimates, we construct time series of daily stablecoin flows from January 1st 2024 to December 31st 2024. For this analysis, we restrict the sample to weekday data only, mitigating mechanical correlation arising from the cylicality of significantly reduced flows on weekends, which is present in both datasets. We begin with gross flows. First, we document a very large, positive correlation between the outflows of the

[27] We provide a comparison of gross flows in Figure 20 in the Appendix.

respective categories in both datasets, ranging from 0.78 to 0.96 (cf. Table 10).[28] Noticeably, correlations between the indirect categories of both datasets are the lowest.

| Region | Direct | Indirect | Direct + Indirect |
|---|---|---|---|
| Africa and Middle East | 0.88 | 0.79 | 0.91 |
| Asia and Pacific | 0.89 | 0.85 | 0.94 |
| Europe | 0.90 | 0.89 | 0.96 |
| Latin America and Caribbean | 0.86 | 0.78 | 0.87 |
| North America | 0.91 | 0.88 | 0.95 |

Table 10: Correlation of Outflows between the Respective Categories in Our Dataset and Chainalysis

Second, we document that even though the correlation between gross flows is very high, a more nuanced picture emerges for net flows. Similar to the pattern we highlighted regarding quantities, there is a high, positive correlation between the direct category in Chainalysis and the direct category in our dataset. However, the indirect category in the Chainalysis dataset exhibits large negative (e.g., for Latin America and Caribbean) to negligible (e.g., for Asia and Pacific) correlation with the indirect category in our dataset, as well as the direct categories in our dataset and the Chainalysis dataset (cf. Table 11). This highlights that the indirect category in the Chainalysis dataset not only disagrees with the direction of net flows of our dataset (and the direct category in Chainalysis), but also with the dynamics of the time series.

| Region | Direct | Indirect | Direct + Indirect | Chainalysis Direct vs Indirect |
|---|---|---|---|---|
| Africa and Middle East | 0.27 | -0.18 | 0.10 | -0.42 |
| Asia and Pacific | 0.75 | 0.01 | 0.50 | -0.13 |
| Europe | 0.22 | -0.15 | 0.35 | -0.27 |
| Latin America and Caribbean | 0.60 | -0.38 | 0.09 | -0.47 |
| North America | 0.80 | -0.22 | 0.53 | -0.39 |

Table 11: Correlation of Net Flows between the Respective Categories in Our Dataset and Chainalysis

Finally, we comment on the correlations between inflows and outflows of the regions within each dataset, as previous research (e.g., Cardozo et al. (2024), Cerutti et al. (2024)) has noted and very high inflow-outflow correlations up to 99% in the Chainalysis dataset.

[28]We report the table for the correlation between inflows, which is quantitatively very similar, in Table 16 in the Appendix.

While we relegate the details to Table 17 in the Appendix, we remark that we find similarly large inflow-outflow correlations to Chainalysis, indicating that they might be a salient feature of the underlying data. Interestingly, calculating inflow-outflow correlations in our dataset purely between self-custodial wallets—which is not possible in the Chainalysis dataset—we find that inflow-outflow correlations are significantly lower, although still high in absolute terms. This hints that high inflow-outflow correlations are particularly salient for flows involving CEXs, and less so for flows involving self-custodial wallets.

# 9 Conclusion

Contrary to prevailing misconceptions, we find that measuring international crypto asset flows, while complex, is not impossible. We develop a novel methodology to estimate the geographic allocation of crypto wallets and employ this approach to quantify international stablecoin flows. We determine that stablecoin flows in 2024 total $2 trillion, the majority of which are international. In absolute terms, we observe the highest volumes in the Asia and Pacific region and North America, whereas we find the lowest volumes in Africa and the Middle East, alongside Latin America and the Caribbean. However, relative to GDP, we find the volumes in these regions to be the most substantial. We establish a correlation between net stablecoin inflows into regions and the relative weakness of domestic currencies against the U.S. dollar, either suggesting that stablecoins serve as a mechanism to fulfill global demand for dollar-based assets for people that seek a hedge against currency depreciation, or that stablecoin flows could possibly be sizable enough to drive exchange rate dynamics. Furthermore, we present evidence of the interconnection between stablecoins and the banking system, highlighting disruptions in stablecoin flows precipitated by the banking crisis of March 2023. We believe that our methodology facilitates a wide range of prospective applications for future research, including the derivation of more granular country-level estimates, the assessment of the geographic distribution of the stock of crypto assets in addition to flows, and the examination of the geographic patterns of decentralized finance application usage.

# References

Alnasaa, M., N. Gueorguiev, J. Honda, E. Imamoglu, P. Mauro, K. Primus, and D. Rozhkov (2022): "Crypto-assets, corruption, and capital controls: Cross-country correlations," *Economics Letters*, 215, 110492.

Aramonte, S., W. Huang, and A. Schrimpf (2022): "Tracing the Footprint of Cryptoization in Emerging Market Economies," *BIS Quarterly Review, Basel, March*.

Arbalik, M., N. Apergis, M. Zeren, and . Sarigül (2021): "International Capital Flows and the Cryptocurrency Effect," *The Journal of Financial Researches and Studies*, 13, 16–35.

Athey, S., I. Parashkevov, V. Sarukkai, and J. Xia (2016): "Bitcoin pricing, adoption, and usage: Theory and evidence," .

Auer, R., U. Lewrick, and J. Paulick (2025): "DeFiying gravity? An empirical analysis of cross-border Bitcoin, Ether and stablecoin flows," *Bank for International Settlements*.

Bank for International Settlements (2023): "BIS Annual Economic Report 2023," .

Calvo, G. A. and C. M. Reinhart (2002): "Fear of Floating," *The Quarterly Journal of Economics*, 117, 379–408.

Cardozo, P., A. Fernández, J. Jiang, and F. Rojas (2024): "On Cross-Border Crypto Flows: Measurement Drivers and Policy Implications," *IMF Working Papers*.

Cerutti, E., J. Chen, and M. Hengge (2024): *A Primer on Bitcoin Cross-Border Flows: Measurement and Drivers*, International Monetary Fund.

Cong, L. W., P. Ghosh, J. Li, and Q. Ruan (2023): "Inflation Expectation and Cryptocurrency Investment," *IIM Bangalore Research Papers*.

Cruz, W. H., J. Xu, P. Tasca, and C. Campajola (2024): "No Questions Asked: Effects of Transparency on Stablecoin Liquidity During the Collapse of Silicon Valley Bank," *arXiv preprint arXiv:2407.11716*.

European Union (2023): "Regulation (EU) 2023/1113 on Information Accompanying Transfers of Funds and Certain Crypto-assets," .

Financial Action Task Force (2023): "Virtual Assets: Targeted Update on Implementation of the FATF Standards on VAs and VASPs," .

Financial Stability Board (2023): "High-level Recommendations for the Regulation, Supervision and Oversight of Crypto-Asset Activities and Markets," .

G7 (2023): "G7 Finance Ministers and Central Bank Governors' Statement on Crypto Assets," .

Gopinath, G. and J. C. Stein (2021): "Banking, Trade, and the Making of a Dominant Currency," *The Quarterly Journal of Economics*, 136, 783–830.

Guillame-Bert, M., S. Bruch, R. Stotz, and J. Pfeifer (2023): "Yggdrasil Decision Forests: A Fast and Extensible Decision Forests Library," in *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, KDD 2023, Long Beach, CA, USA,* *August 6-10, 2023*, 4068–4077.

Hastie, T., R. Tibshirani, and J. Friedman (2017): *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer.

He, H. and E. A. Garcia (2009): "Learning from imbalanced data," *IEEE Transactions on knowledge and data engineering*, 21, 1263–1284.

Hu, Maggie R., L., A. D., and T. J. Putniņš (2021): "Evading Capital Controls via Cryptocurrencies: Evidence from the Blockchain," *SSRN Electronic Journal*.

International Monetary Fund (2023): "Elements of Effective Policies for Crypto Assets," .

King, G. and L. Zeng (2001): "Logistic regression in rare events data," *Political analysis*, 9, 137–163.

Makarov, I. and A. Schoar (2021): "Blockchain Analysis of the Bitcoin Market," .

——— (2022): "Cryptocurrencies and Decentralized Finance (DeFi)," *Brookings Papers on Economic Activity*, 2022, 141–215.

Medvedev, E. and the D5 team (2018): "Ethereum ETL," .

Meiklejohn, S., M. Pomarole, G. Jordan, K. Levchenko, D. McCoy, G. M. Voelker, and S. Savage (2013): "A fistful of bitcoins: characterizing payments among men with no names," in *Proceedings of the 2013 conference on Internet measurement conference*, 127–140.

United States Department of the Treasury (2023): "Treasury Illicit Finance Risk Assessment of Decentralized Finance," .

von Luckner, C. G., R. Koepke, and S. Sgherri (2024): *Crypto as a Marketplace for Capital Flight*, International Monetary Fund.

von Luckner, C. G., C. M. Reinhart, and K. Rogoff (2023): "Decrypting New Age International Capital Flows," *Journal of Monetary Economics*, 138, 104–122.

World Bank (2024): "Remittance Prices Worldwide Quartlers," .

# A Regional Centralized Exchanges

| Region | Exchange | Rationale |
|---|---|---|
| Africa and Middle East | Altcointrader | Headquartered in South Africa |
| Africa and Middle East | Arzpaya.com | Focused on Iran region |
| Africa and Middle East | Artis Turba Exchange | Headquartered in Middle East |
| Africa and Middle East | Bit2c | Headquartered in Israeli region |
| Africa and Middle East | Bitoasis | Regulated in Middle East |
| Africa and Middle East | Luno | Operating in South Africa |
| Africa and Middle East | Nobitex | Licensed in Iran region |
| Africa and Middle East | Valr | South Africa based service |
| Asia and Pacific | Bitbank | Mainly Japan market focus |
| Asia and Pacific | Bitkub | Licensed in Thailand region |
| Asia and Pacific | Bithumb | Operating in South Korea |
| Asia and Pacific | Coindcx | Headquartered in India region |
| Asia and Pacific | Coincheck | Based in Japan market |
| Asia and Pacific | Coinhako | Licensed in Singapore region |
| Asia and Pacific | Coinone | South Korea market focus |
| Asia and Pacific | Coins.ph | Operating in Philippine market |
| Asia and Pacific | Gdac | South Korea service focus |
| Asia and Pacific | GMO Coin | Operating in Japanese market |
| Asia and Pacific | Gopax | Licensed in South Korea |
| Asia and Pacific | Indodax | Focus on Indonesian market |
| Asia and Pacific | Korbit | South Korea regulatory focus |
| Asia and Pacific | Maicoin | Based in Taiwan Province of China market |
| Asia and Pacific | Tokocrypto | Indonesian market primary focus |
| Asia and Pacific | Upbit | Primary market: South Korea |
| Asia and Pacific | Wazirx | Operating in Indian market |
| Europe | Anycoin Direct | Licensed in Netherlands region |
| Europe | Bitpanda | Based in Austria region |
| Europe | Bitvavo | Operating in Netherlands market |
| Europe | Btcturk | Licensed in Turkey region |
| Europe | Coinmetro | Headquartered in Malta region |
| Europe | Exmo | Focus on Eastern Europe |
| Europe | Firi | Based in Spain region |
| Europe | Norwegian Block Exchange | Headquartered in Norway region |
| Europe | Paribu | Operating in Turkey market |
| Europe | Swissborg | Licensed in Switzerland region |
| Latin America and Caribbean | Bitso | Based in Mexico region |
| Latin America and Caribbean | Brasil Bitcoin | Operating in Brazil market |
| Latin America and Caribbean | C-Patex | Focus on Latin America |
| Latin America and Caribbean | Lemon Cash | Serving Latin America clientele |
| Latin America and Caribbean | Mercado Bitcoin | Licensed in Brazil region |
| Latin America and Caribbean | Orionx | Brazil market regulatory focus |
| Latin America and Caribbean | Panda Exchange | Primarily Latin America focus |
| North America | Binance US | Based in United States |
| North America | Bitbuy | Headquartered in Canada region |
| North America | Coinsquare | Licensed in Canadian market |
| North America | Netcoins | Operating in Canadian market |
| North America | Quadrigacx | Based in Canada region |
| North America | Shakepay | Licensed in Canada region |

Table 12: Exchanges by Region with Brief Rationale

# B Daylight Savings Time Graphs

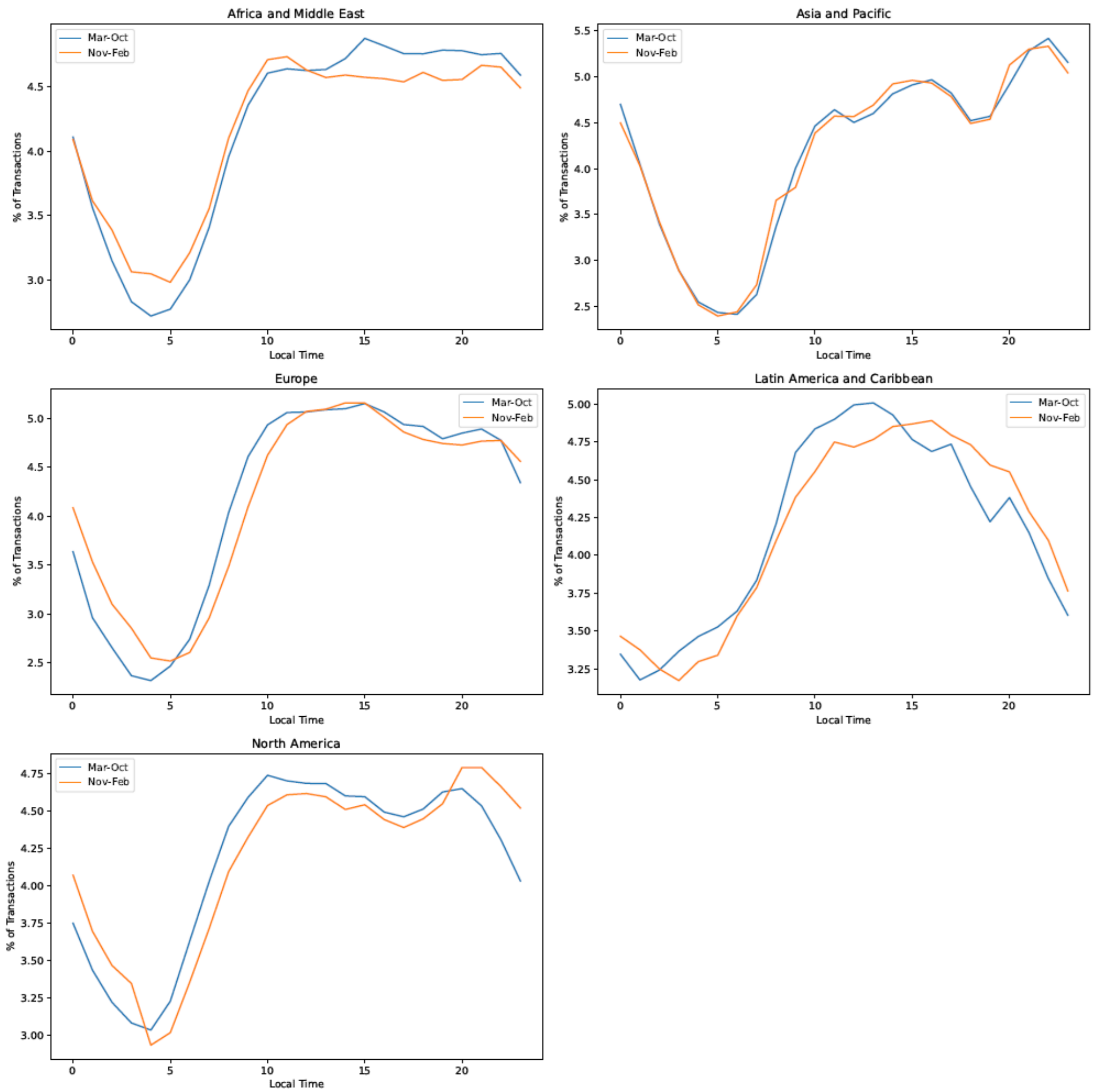



Figure 15: Figures Depicting Activity During DST and no-DST Months.

# C Identifying Additional Centralized Exchange Wallets

We briefly describe how centralized exchanges structure their wallets. There are 4 different types: hot wallets, cold wallets, deposit wallets and gas supplier wallets. Hot wallets are wallets that hold pooled customer funds and are used to send out funds from the exchange to other wallets in on-chain transactions. Cold wallets are wallets that exchange use for long term storage of funds, are not relevant for our analysis. Last, there are gas supplier wallets that provide gas to other exchange wallets when needed.[29] These addresses are present in the dataset that we get from Dune.

Deposit wallets are wallets that are specifically created for customers when they want to deposit funds into the exchange. These wallets are absent from the Dune dataset. Typically, these addresses immediately forward deposits into the exchange's hot wallets and do not engage in any other activities. We identify a wallet as a deposit wallet belong to a particular exchange if it (1) only forward funds to hot wallets of an exchange or (2) has receives gas from a gas supplier wallet of an exchange.

# D Model Evaluation

[29]Gas measures the computational effort needed to process on-chain transactions. Users must pay for gas used, typically with the native protocol crypto asset, which for the blockchains in our sample is Ether, except on Binance Smart Chain which uses Binance Coin.

**accuracy:** 0.697201
**PR-AUC: 'Asia and Pacific' vs others:** 0.853099
**PR-AUC: '{Africa and Middle East, Europe}' vs others:** 0.760129
**PR-AUC: '{Latin America and Caribbean, North America}' vs others:** 0.668412
**num examples:** 32259

**AUC: 'Asia and Pacific' vs others:** 0.849414
**AUC: '{Africa and Middle East, Europe}' vs others:** 0.831646
**AUC: '{Latin America and Caribbean, North America}' vs others:** 0.865555
**loss:** 0.713695
**num examples (weighted):** 32259

**Confusion matrix**

| Label \ Pred | Asia and Pacific | {Africa and Middle East, Europe} | {Latin America and Caribbean, North America} |
|---|---|---|---|
| Asia and Pacific | 10857 | 2119 | 517 |
| {Africa and Middle East, Europe} | 3245 | 8833 | 870 |
| {Latin America and Caribbean, North America} | 1445 | 1572 | 2801 |

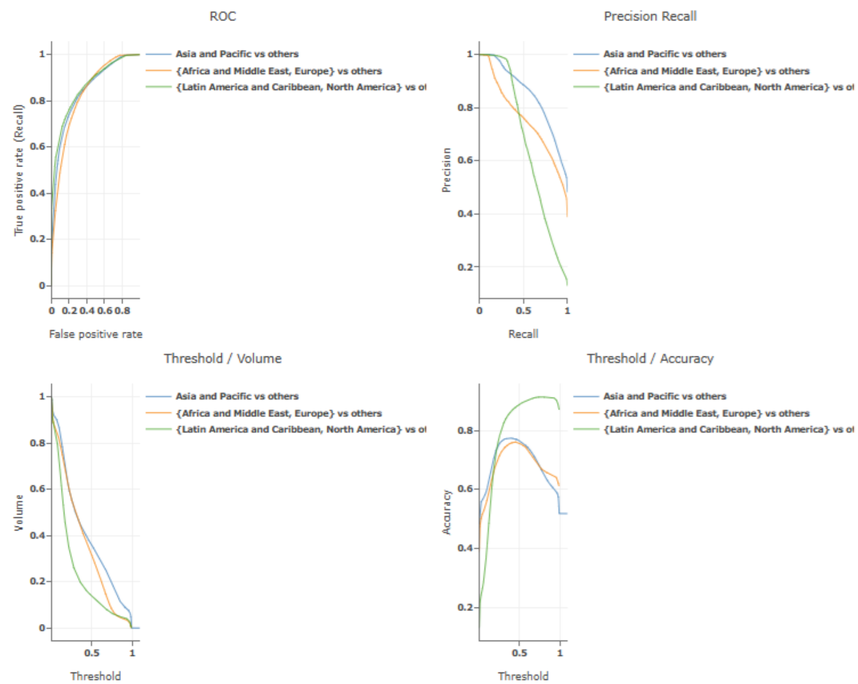



Figure 16: Model Evaluation for Classification of Asia and Pacific vs {Africa and Middle East, Europe} vs. {Latin America and Caribbean, North America}

**accuracy:** 0.844556 **AUC: 'North America' vs others:** 0.894665
**PR-AUC: 'North America' vs others:** 0.979059 **loss:** 0.334371
**num examples:** 4188 **num examples (weighted):** 4188

**Confusion matrix**

| Label \ Pred | Latin America and Caribbean | North America |
|---|---|---|
| **Latin America and Caribbean** | 428 | 475 |
| **North America** | 176 | 3109 |

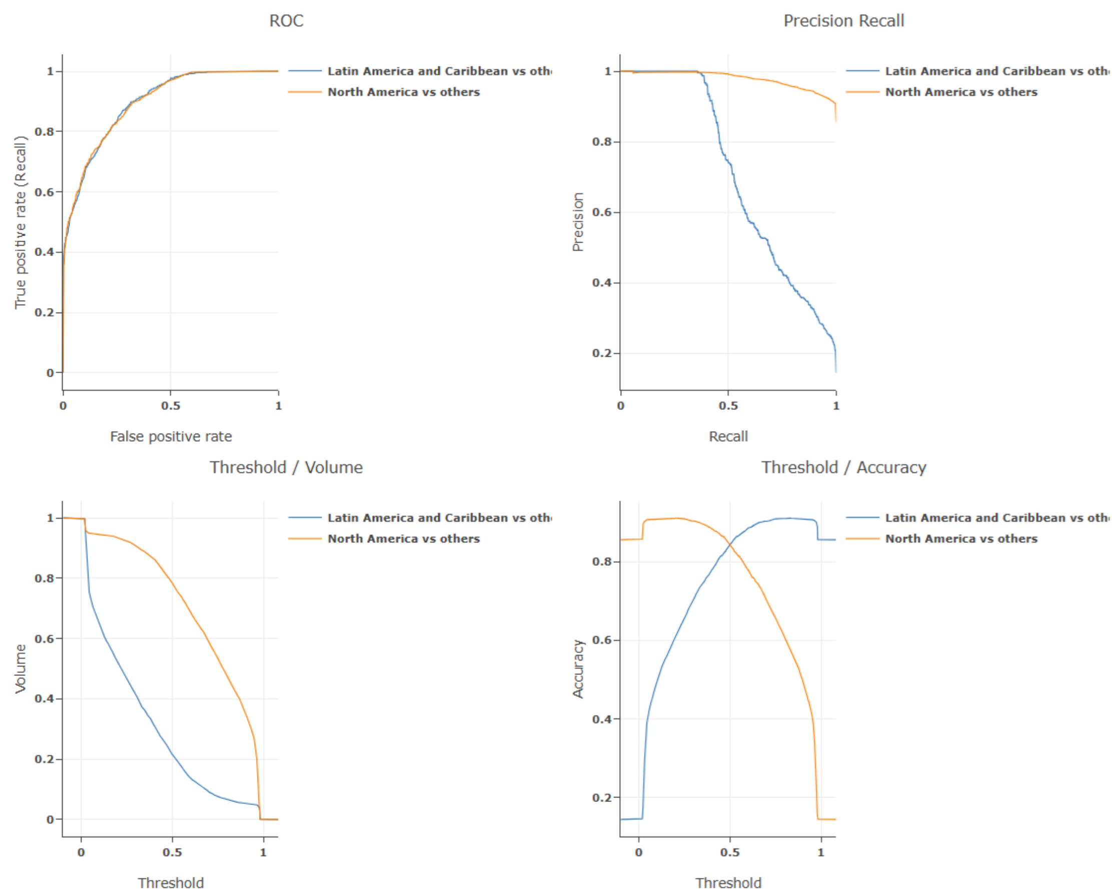



Figure 17: Model Evaluation to Distinguish Latin America and Caribbean from North America

**accuracy:** 0.782098 **AUC: 'Europe' vs others:** 0.805675
**PR-AUC: 'Europe' vs others:** 0.930036 **loss:** 0.476119
**num examples:** 12524 **num examples (weighted):** 12524

**Confusion matrix**

| Label \ Pred | Africa and Middle East | Europe |
|---|---|---|
| **Africa and Middle East** | 1641 | 1808 |
| **Europe** | 921 | 8154 |

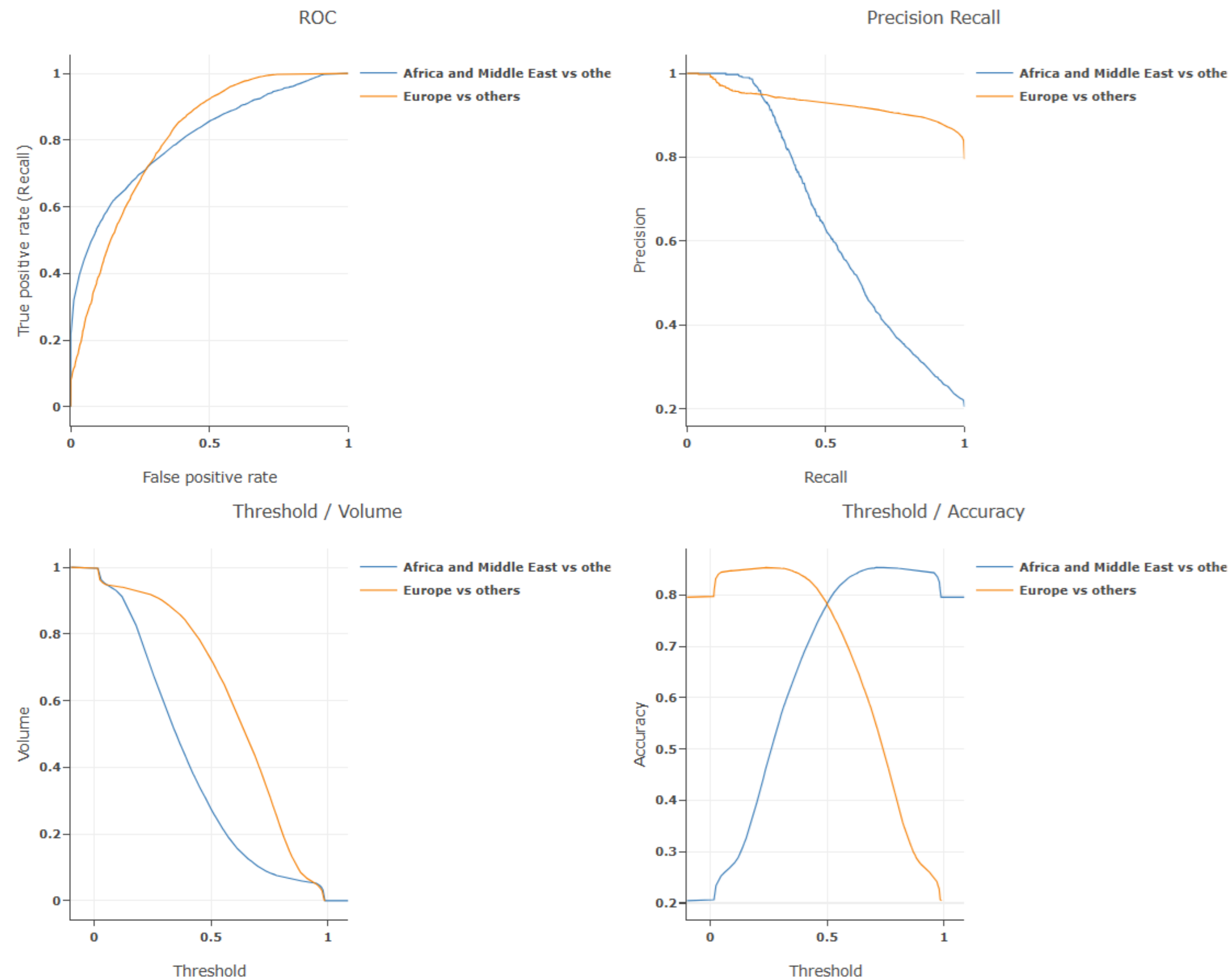



Figure 18: Model Evaluation to Distinguish Africa and Middle East from Europe

**accuracy:** 0.823239 **AUC: 'non-China' vs others:** 0.911902
**PR-AUC: 'non-China' vs others:** 0.96429 **loss:** 0.392763
**num examples:** 155464 **num examples (weighted):** 155464

**Confusion matrix**

| Label \ Pred | China | non-China |
|---|---|---|
| **China** | 37068 | 20940 |
| **non-China** | 6540 | 90916 |

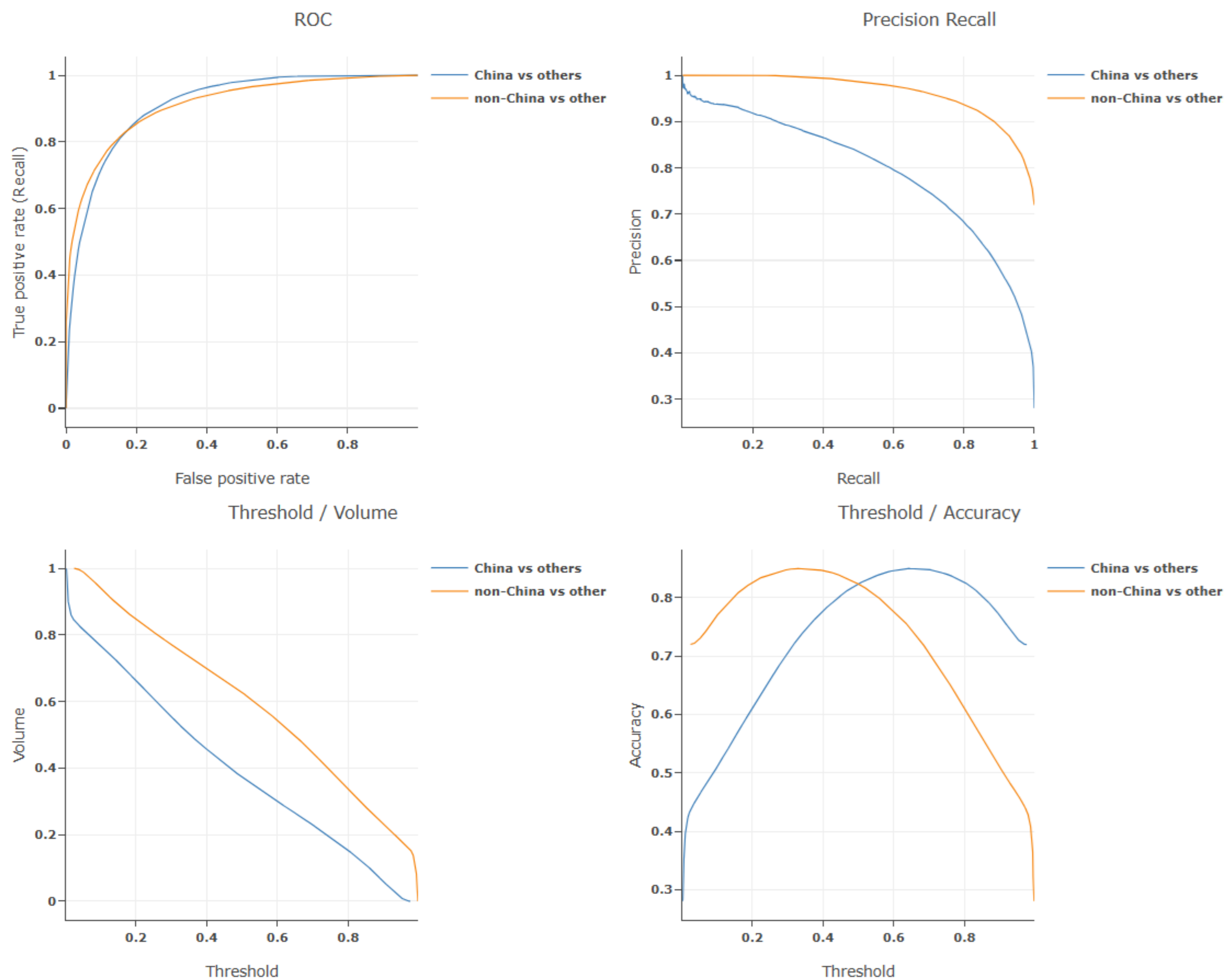



Figure 19: Model Evaluation to Distinguish between China and Asia and Pacific (excluding China)

# E Manually Assigned Wallets

| **Address** | **Region** | **Entity** |
|---|---|---|
| 0x55fe002aeff02f77364de339a1292923a15844b8 | North America | Circle |
| 0xad6eaa735d9df3d7696fd03984379dae02ed8862 | North America | Cumberland |
| 0x87b49a99cbce4a9030e67919b776aa97d538adda | North America | Cumberland |
| 0xf584f8728b874a6a5c7a8d4d387c9aae9172d621 | North America | Jump Trading |
| 0x0548f59fee79f8832c299e01dca5c76f034f558e | North America | Genesis Trading |
| 0xd628f7c481c7dd87f674870bec5d7a311fb1d9a2 | North America | Genesis Trading |
| 0x84d34f4f83a87596cd3fb6887cff8f17bf5a7b83 | North America | Alameda Research |
| 0xe31a9498a22493ab922bc0eb240313a46525ee0a | North America | Alameda Research |
| 0x17d70306956a6a4b4f9319ad9b9de43e98382f5e | North America | Alameda Research |
| 0x2f2be7c998a2abcf0caa32d1b7da714ea0a0e2d2 | North America | Alameda Research |
| 0x83a127952d26ced22410cb1dbe4bfe2676bc63bd | North America | Alameda Research |
| 0xb560da83a2c351fca35e5ebadba2a82fd525d4c3 | North America | Alameda Research |
| 0x1d77f556ee0dbd8b07a7bd4fa461ad24d35543ba | North America | Alameda Research |
| 0x0ae80df72ad0620b1d34d1ec31fa43415bfe0afc | North America | Alameda Research |
| 0x882a8127d5aee37c82ba1449f28e1252e3ee6620 | North America | Alameda Research |
| 0x82a505ad68bc9a10a96f807df60078ef75bd5e56 | North America | Alameda Research |
| 0x01811f428f03682d43db8d1bbf242dcd05acbe9f | North America | Alameda Research |
| 0x8be32560a42a378d349ba0d69d54b210b31d9efb | North America | Alameda Research |
| 0xa8553cfb14d2321f0cf2cadae36bf2d607a552ed | North America | Alameda Research |
| 0x9ea14a8379152f42d39d24239100ca4546722d92 | North America | Alameda Research |
| 0x75ec94e298dc0e3b00c30955c94edb40049a2a44 | North America | Alameda Research |
| 0x4a9f2de50756c756fad90c3037bf1f39676ff701 | North America | Alameda Research |
| 0x67dce0c45fc2e38812a8602ea6ff7b4eb90c839b | North America | Alameda Research |
| 0xb9bd20ec7b4d24bc115ef24724ad04d851b2b9b0 | North America | Alameda Research |
| 0x0a4d88a90b0b9c53bd2d167fede915ffbe2238fe | North America | Alameda Research |
| 0x7a66dc0da224955e8256d9c289ef345c7cb8d229 | North America | Alameda Research |
| 0x875b7f1d8f1986f369dd08c801ef47f64e8c320a | North America | Alameda Research |
| 0x7cf4ce48bf3b7e3139c25c017978a71b2ba293be | North America | Alameda Research |
| 0x84806f88e475f556883a607e1d9b0c3fe79ef15f | North America | Alameda Research |
| 0x30da8f270a92a2ab076392b4ab72bfaa476ca1d1 | North America | Alameda Research |
| 0xa8cfec07a38c5b3fa0b5ae7fe1f71412ced385fa | North America | Alameda Research |
| 0x3507a4978e0e83315d20df86ac0b976c0e40ccb | North America | Alameda Research |
| 0x83a127952d266a6ea306c40ac62a4a70668fe3bd | North America | Alameda Research |
| 0xdbf5e9c5206d0db70a90108bf936da60221dc080 | Europe | Wintermute |
| 0x8aceab8167c80cb8b3de7fa6228b889bb1130ee8 | Europe | Celsius |
| 0x4862733b5fddfd35f35ea8ccf08f5045e57388b3 | Asia and Pacific | Three Arrows Capital |
| 0x3ddfa8ec3052539b6c9549f12cea2c295cff5296 | Asia and Pacific | Justin Sun |

Table 13: Blockchain addresses manually assigned to regions

# F Regression Tables

## F.1 Robustness Checks for Link Between Stablecoins and Exchange Rates

| | Global | China |
|---|---|---|
| VIX | 0.026 | 0.036 |
| | (0.065) | (0.049) |
| Broad Dollar | 0.122 | |
| | (0.078) | |
| USD/CNY | | 0.182** |
| | | (0.081) |
| Crypto F&G | 0.042 | 0.003 |
| | (0.038) | (0.044) |
| $\text{Flows}_{t-1}$ | 0.334*** | 0.392*** |
| | (0.041) | (0.076) |
| Region × Quarter FE | ✓ | |
| Quarter FE | | ✓ |
| Weekend FE | ✓ | ✓ |
| Observations | 4·1095 | 1095 |
| R-squared | 0.399 | 0.324 |
| F-statistic | 58.543 | 32.212 |

Standard errors in parentheses

$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

Region: Standard errors clustered at the time level.

China: Robust standard errors.

Table 14: Panel Regression for Drivers of Net Stablecoin Flows

## F.2 Differences-in-differences estimates for March 2023 Banking Crisis

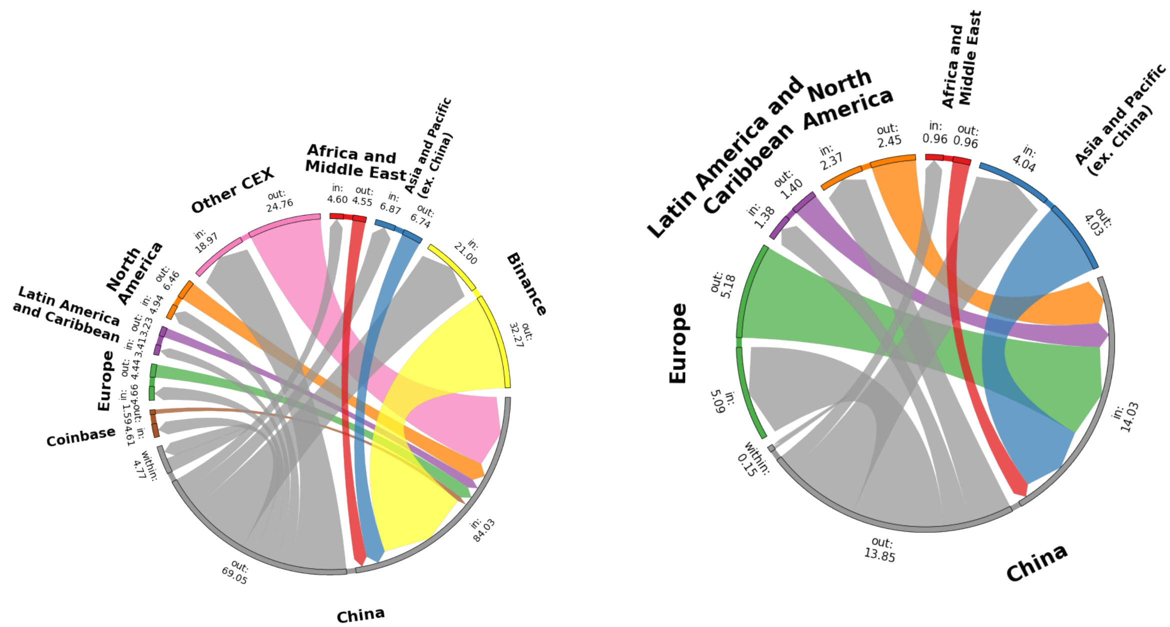
| | |
|---|---|
| Treated × $\tau = -5$ | -0.381 |
| | (0.359) |
| Treated × $\tau = -4$ | 0.257 |
| | (0.365) |
| Treated × $\tau = -3$ | 0.294 |
| | (0.417) |
| Treated × $\tau = -2$ | 0.477 |
| | (0.343) |
| Treated × $\tau = 1$ | -9.955*** |
| | (1.510) |
| Treated × $\tau = 2$ | -4.540*** |
| | (0.720) |
| Treated × $\tau = 3$ | -1.290*** |
| | (0.320) |
| Treated × $\tau = 4$ | -0.947** |
| | (0.352) |
| Treated × $\tau = 5$ | -0.826** |
| | (0.377) |
| Observations | 50 |
| R-squared | 0.723 |
| F-statistic | 7.813 |
| Fixed Effects: Region, Time | |

Standard errors in parentheses
$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$
Robust standard errors.

Table 15: Difference-in-Differences Regression for the Effect of the March 2023 Banking Crisis on Stablecoin Flows

# G Supplementary Material for Comparison with the Chainalysis Dataset

## G.1 Comparison of China Gross Flows Between Datasets




(a) China gross flows in our data

(b) China gross flows in Chainalysis (Direct + Indirect)

Figure 20: Comparison of gross stablecoin flows involving China between both datasets.

## G.2 Comparison of Correlations Between Datasets

| Region | Direct | Indirect | Direct + Indirect |
|---|---|---|---|
| Africa and Middle East | 0.89 | 0.83 | 0.93 |
| Asia and Pacific | 0.93 | 0.87 | 0.95 |
| Europe | 0.92 | 0.90 | 0.96 |
| Latin America and Caribbean | 0.89 | 0.84 | 0.91 |
| North America | 0.81 | 0.83 | 0.94 |

Table 16: Correlation of Inflows between the Respective Categories in Our Dataset and Chainalysis

| | Our Dataset | | | | Chainalysis | | |
|---|---|---|---|---|---|---|---|
| Region | Direct | Indirect | Total | Self-custodial | Direct | Indirect | Total |
| Africa and Middle East | 0.99 | 0.98 | 0.99 | 0.89 | 0.96 | 0.96 | 0.98 |
| Asia and Pacific | 0.94 | 0.98 | 0.98 | 0.88 | 0.97 | 0.98 | 0.99 |
| Europe | 0.99 | 0.98 | 0.99 | 0.84 | 0.99 | 0.99 | 1.00 |
| Latin America and Caribbean | 0.99 | 0.99 | 0.99 | 0.91 | 0.97 | 0.97 | 0.98 |
| North America | 0.94 | 0.93 | 0.95 | 0.66 | 0.94 | 0.93 | 0.96 |

Table 17: Inflow-Outflow Correlations by Categories for Both Datasets
Note: We exclude within-region flows from the data, as these represent both inflows and outflows, introducing mechanical inflow-outflow correlation.

PUBLICATIONS

**Decrypting Crypto: How to Estimate International Stablecoin Flows**

Working Paper No. WP/2025/141