Network Topologies of Intermediaries in the Offshore World

Oscar Julius Adserballe
Student ID: S160855

Copenhagen Business School

Supervisor: Rasmus Corlin Christensen

May 12, 2025

Contents

Abstract **1**

1 Introduction & Motivation **2**
1.1 Introduction . . . 2
1.2 Tax Avoidance at the Top of the Income Distribution . . . 3
1.3 Limitations of Traditional Demand-Side Models . . . 3
1.4 The Supply-side: Intermediaries as Gatekeepers . . . 4
1.5 Research Gap: Understanding the *Network Structure* to Inform Intermediary Regulation . . . 6
1.6 RQ: What role do offshore intermediaries play in networks of high-end tax avoidance? . . . 7
1.7 Roadmap of the Thesis. . . . 7

2 Theory **8**
2.1 A Note on Philosophy of Science and Methodological Approach . . . 8
2.2 Conceptual foundations . . . 8
2.2.1 Global Wealth Chains and the Role of Intermediaries . . . 8
2.2.2 Weaponised Interdependence . . . 10
2.2.3 Network Theory as a Lens for Understanding Illicit networks . . . 11
2.2.4 A Typology of Intermediaries and Their Role . . . 12
2.2.5 Secrecy Strategies: Financial Instruments and Legal Innovations . 13
2.3 Propositions . . . 14

3 Data and Methodology **15**
3.1 The ICIJ Offshore Leaks Database . . . 15
3.2 External Data Sources . . . 16
3.3 Using Agentic AI to Scrape Data on Intermediaries. . . . 17
3.4 Use of LLMs . . . 23
3.5 Use of LLMs in the Broader Paper . . . 23

4 Empirical Analysis **24**

5 Discussion **34**

6 Conclusion **35**

A Appendix **36**

List of Figures

3.1 Agent Setup for Intermediary Classification 18

4.1 Entity Incorporations Over Time by Source Investigation. 24
4.2 Illustration of the Power Law Distribution of Intermediary Connections (Degree). 25
4.3 Coverage of Intermediary Connections Explained by the Enriched Top Intermediaries. 25
4.4 Geographic Distribution of Intermediaries (Top Locations). 26
4.5 Coverage of Intermediaries by Top Geographic Locations. 26
4.6 Geographic Distribution of Entities (Top Jurisdictions/Locations). 27
4.7 Coverage of Entities by Top Geographic Locations. 27
4.8 Geographic Distribution of Officers (Top Locations). 28
4.9 Coverage of Officers by Top Geographic Locations. 28
4.10 Comparison of Intermediary Classification Distributions (Random Sample vs. Top 5%). 29
4.11 Cumulative Distribution Function (CDF) of Intermediary Degree by Classification. 29
4.12 Boxplot of Intermediary Degree by Classification. 30
4.13 Distribution of Intermediary Classifications across Top Countries. 30
4.14 Temporal Evolution of the Intermediary-Entity Network (Cyprus Case Study). Shows connections between intermediary types (based in Cyprus) and the locations of entities they manage at different time points. 33

List of Tables

Abstract

Intermediaries form the crucial links enabling tax haven use, contributing significantly to global tax avoidance and inequality. While analyses often focus on demand-side factors, this thesis challenges such perspectives by asserting the critical importance of the supply-side network structure and intermediary agency for understanding and regulating offshore finance. Extending the network analysis of Chang et al. (2023) and drawing on Harrington's (2016) micro-sociological evidence, I analyze International Consortium of Investigative Journalists (ICIJ) leak data. Although ICIJ data has limitations for estimating the overall scale of avoidance, it permits robust generalization about intermediary roles within these complex networks. A novel agentic method is employed, enriching ICIJ data by incorporating publicly available online information about intermediaries' professional roles and affiliations.

This thesis presents four key propositions: 1) The overall network exhibits structural vulnerabilities concentrated around central intermediaries. 2) Intermediaries often display cultural or national specificity, catering to distinct clienteles. 3) Different intermediary types occupy distinct network positions and vary in systemic importance, measurable via network centrality. 4) Network structures are dynamic, adapting in response to regulation and financial innovation. This analysis provides critical insights into the architecture and potential regulatory chokepoints within tax haven networks, viewed through the lens of intermediary action.

Introduction & Motivation

1.1 Introduction

The central claim advanced throughout this thesis concerns the critical relevance of examining supply-side dynamics within the offshore financial system. Specifically, it argues that the role of intermediaries – the professional enablers and facilitators of offshore activity – is an incredibly relevant factor. The function and influence of the supply side – encompassing the specialized intermediaries and the specific services offered by various jurisdictions that actively enable and shape offshore activity – remains comparatively under-explored from an empirical standpoint. Building upon recent scholarship that increasingly highlights these supply dynamics (e.g., Laffitte 2024; Alstadsæter et al. 2019), this thesis seeks to extend and generalize insights from qualitative work, such as Harrington's (2016) study of wealth managers, through a quantitative analysis drawing upon the extensive data revealed by the ICIJ leaks.

Primary literature this is building on (contextualising interest in the topic):

- Interest spurred on this by an interest in optimal taxation regimes esp. Saez (2002), and the work of Zucman & Saez (2019) on the optimal taxation of wealth.
- Overall approach from neoclassical public finance and economics. Lectures from Zucman's overviews of tax evasion and avoidance in the modern economic literature has been the primary source. https://gabriel-zucman.eu/publicecon/
- Niche within Political Sociology through Brooke Harrington (2016)'s book and the method's employed in her ethnography of wealth managers. Likewise the tentative work in Chang et al. (2023a and 2023b) on network structure. However, for the latter, they concentrate more on demand-strategies rather than the more interesting supply-side strategies that are the focus of thisthesis.

1.2 Tax Avoidance at the Top of the Income Distribution

While considerable progress has arguably been made in curbing outright tax evasion, tax avoidance remains a substantial challenge, a point emphasized by commentators such as Stiglitz (cited in Alstadsæter et al., 2024). It introduces several clear inefficiencies into the economic system, including the generation of a distinct class and socially unoptimal rents accruing to the intermediaries who facilitate such schemes, the potential for poor allocation of resources as investment decisions are distorted by spurious incentives, and, beyond these economic inefficiencies, a range of normative concerns regarding fairness and the integrity of the tax system that inevitably accompany widespread tax avoidance.

A literature that has grown very prominent in the past two decades or so in A crucial distinction often highlighted is between income and wealth inequality. Income inequality can be somewhat ephemeral in nature; high-earners in one year may retire or experience income fluctuations in the next. Wealth, in contrast, tends to be more permanent, potentially distorting social outcomes over non-transient periods in a more meaningful way. Inordinate wealth accumulation (e.g. Harrington, 2016) has distorted social mobility (as explored in the work of Chetty) and been a key driver of overall inequality trends (e.g. Piketty's main body of work).

With that said, from a (narrow and purely economic) point of view, whether tax avoidance quantifying is actually bad is unclear, so the normative desirability of it at aggregate is still in question. The precise behavioral effects of tax evasion and avoidance on incentives – such as the incentives to work, save, or invest – is not as clear as, for example, studying the effects of tax incentives on MNCs (where it seems generally negative, e.g. Puerto Rico tax credit study from Serrato, 2018; also Garrett & Serrato, 2019). A key complicating factor is the role of expectations; an individual's behavior is likely highly dependent on their expectation of being able to successfully evade or avoid taxes in the future.

1.3 Limitations of Traditional Demand-Side Models

Traditionally, tax evasion and avoidance has been studied from the demand-side. The seminal Allingham-Sandmo (1972) good at explaining tax evasion decisions of the vast majority of the income distribution (Alstadsæter et al. 2019) performs poorly at the top of the distribution (ibid.) the Allingham-Sandmo (1972) model, provides a powerful and often empirically supported framework for understanding tax evasion decisions for the majority of taxpayers. This standard model typically portrays evasion as a individual and rational gamble, where individuals weigh the expected benefits of non-compliance against

the probability of detection and the severity of potential penalties (see also Yitzaki & Slemrod). However, under standard assumptions about risk aversion and the structure of penalties and audit probabilities, the model often predicts that wealthier individuals, facing potentially higher stakes and scrutiny, should be less inclined to evade taxes. Yet, empirical evidence, particularly from studies leveraging leaked data (e.g., Alstadsæter et al. 2019), suggests the opposite: offshore tax evasion appears highly concentrated among the ultra-wealthy. The comparative statics do not hold here.

Furthermore, traditional methods for empirically studying tax compliance, such as random audit studies (e.g., Kleven et al. 2011), also face limitations in capturing the full picture of high-end evasion. As highlighted by Alstadsæter et al. (2019), while random audits are invaluable for understanding compliance behavior regarding income streams typically subject to third-party reporting or easily verifiable through standard audits, they often fail to detect the sophisticated, cross-border evasion strategies frequently utilized by the wealthiest segment. Complex offshore structures, shell corporations, and opaque trust arrangements often fall outside the scope of conventional audit procedures, rendering this form of evasion largely invisible to standard demand-side enforcement tools.

This points towards a dynamic of a game of cat and mouse. Demand-side enforcement mechanisms, predicated on detecting and penalizing individual non-compliance, struggle to keep pace with the evolving and increasingly complex strategies developed to obscure wealth and income, often with the assistance of specialized intermediaries. Consequently, relying solely on demand-side models and traditional enforcement metrics provides an incomplete, and potentially misleading, understanding of the phenomenon, especially concerning the significant evasion occurring at the top of the distribution. This underscores the necessity of incorporating supply-side factors and network structures to actually understand these mechanisms enabling tax avoidance at the top of the income distribution.

1.4 The Supply-side: Intermediaries as Gatekeepers

To fully grasp the dynamics of offshore tax evasion and avoidance, it is crucial to clarify what constitutes the "supply-side" (used more-so metaphorically than stringently) in this context. Here, the supply-side refers specifically to the ecosystem of professional intermediaries – such as law firms, banks, trust companies, and specialized advisors – as well as the jurisdictions that provide the legal and regulatory frameworks enabling offshore financial activities. The central argument advanced in this thesis, building on insights from models like Alstadsæter et al. (2019) and qualitative work such as Harrington (2016), is that this supply-side dimension is far more relevant to scrutinize than often acknowledged, potentially offering more effective avenues for understanding and potentially curbing offshore practices compared to a sole focus on demand-side factors.

A primary reason for emphasizing the supply side relates to the concept of elasticity. It is argued here that the elasticity of supply of intermediaries is considerably higher, and therefore potentially more responsive to policy interventions, compared to the elasticity of demand from clients seeking offshore services. Several factors underpin this view:

First, the incentives structuring the behavior of intermediaries are arguably much more sensitive to changes in the regulatory or reputational environment. For these professionals and firms, the provision of offshore services is not merely an option but often a core component of their business model and career trajectory. Their professional existence and profitability are directly dependent on their continued ability to offer these specific services effectively and discreetly. Consequently, factors that threaten this ability – such as increased regulatory scrutiny, heightened enforcement risk, or significant reputational damage – can have a pronounced impact on their willingness and capacity to supply these services. In contrast, the demand for tax minimization or evasion among potential clients, driven by factors like high tax rates or a desire for secrecy, can be seen as a relatively persistent force. While demand might fluctuate, the fundamental desire among some wealthy individuals and corporations to reduce tax burdens is likely to remain, making demand potentially less elastic to targeted interventions than the specialized supply of enabling services.

Second, the micro-sociological account provided by Harrington (2016) and Hoang (2022) offers compelling reasons why intermediaries are so central. Her ethnographic work illuminates the deeply personal, trust-based relationships that often form between wealth managers and their elite clients. These relationships, built over time and predicated on discretion and expertise, are difficult to replace. Clients rely heavily on their chosen intermediaries not just for technical execution but also for navigating the complexities and risks of the offshore world. The non-substitutable nature of these trust-based relationships means that disrupting the intermediary side can significantly impact clients' access to and ability to maintain offshore structures, further highlighting the critical role of the supply-side actors.

Third, the structure of the market itself points towards the strategic importance of intermediaries. There often exists a many-to-one relationship between clients and intermediaries; that is, a relatively small number of specialized intermediary firms or key professionals service a large number of clients seeking offshore solutions. This concentration means that the intermediary sector represents a point of leverage. Regulatory actions or enforcement efforts focused on these key intermediary players could potentially have a cascading effect, impacting a wide network of clients far more efficiently than attempting to identify and pursue each individual client separately. This structural feature makes the intermediary supply-side particularly vulnerable, and thus relevant, from a regulatory perspective.

1.5 Research Gap: Understanding the *Network Structure* to Inform Intermediary Regulation

Considerable research, particularly micro-sociological accounts like Harrington's (2016) ethnography, provides rich insights into the dyadic relationships, motivations, and practices of individual wealth managers and their clients. Ethnography, as a methodology, certainly offers a powerful means of accessing and understanding micro-level dynamics that can illuminate macro-level phenomena or "megatrends,"; of "entering in" an otherwise abstract metanarrative (cf. Neely, 2021; Also Chung 2018(check up; misremeber?)) However, generalizing from these detailed qualitative studies to broader systemic patterns has not really been done.

A nascent thread of literature has begun to explore these structural aspects, often spurred by the availability of large-scale leaked data. Work such as Chang et al. (2023), alongside policy-oriented research emerging from bodies like the EU following disclosures such as the Panama Papers (e.g., research from 2017), represents initial steps in this direction. However, this line of inquiry remains limited thus far, often focusing on specific subsets of countries or actors. The analysis of the network structures inherent in the offshore world is still in a highly exploratory phase. Consequently, the potential held within detailed micro-data sources, such as the ICIJ leaks which map connections between entities, officers, and intermediaries on a vast scale, remains largely underexplored in terms of systematic structural analysis.

The work by Chang et al. (2023) on "Secrecy Strategies" provides a pertinent example. While their primary focus was on analyzing the demand strategies employed by global elites, their findings crucially demonstrate that these strategies are shaped by, and interact with, the supply landscape – the available intermediaries, jurisdictions, and the institutional context of the elites' home countries. Their research, therefore, implicitly highlights the importance of the supply structure by showing how it influences demand patterns, effectively linking the two sides of the market through observable strategic choices.

This points towards the specific research gap addressed herein: the need for a more systematic understanding of the network structure of the supply-side itself. While we have compelling accounts of individual intermediary roles and incentives, a comprehensive picture of how these intermediaries connect to each other, to different types of clients, across various jurisdictions, and through specific service offerings – essentially, the topology of the intermediary network – is lacking. Understanding this structure is potentially crucial for designing more effective regulation targeting these key players.

Therefore, the goal within this thesis is to contribute to bridging this gap, primarily through synthesis and systematization. Drawing upon the existing literature, including

the rich ethnographic accounts, the aim here is not necessarily to conduct a novel quantitative network analysis but rather to attempt to codify more generally and quantitatively on some of the more loosely defined observations about intermediaries and their roles. By viewing these observations through the conceptual lens of network structures and positions, the objective is to formulate more general propositions regarding intermediary behavior, influence, and potential vulnerabilities within the broader offshore system.

1.6 RQ: What role do offshore intermediaries play in networks of high-end tax avoidance?

1.7 Roadmap of the Thesis.

Having gone through what motivates the pursuit of this question and situate this thesis, will proceed to the bulk of the paper. First, outline the key concepts and theories I will draw on, then moving on to outline the key propositions this paper will seek to set forth about the role of intermediaries. Then, a brief section will cover the data sources.

Theory

2.1 A Note on Philosophy of Science and Methodological Approach

In line with perspectives advocating for methodological pluralism and the use of qualitative insights for broader theory development (e.g., George & Bennett, 2005), this thesis leverages Harrington's findings for concept formation and hypothesis generation - effectively using her work as a theory-building step. Her work helps define the "intermediary" phenomenon and suggests the importance of factors like trust and expertise, which likely underpin the network structures we observe. This thesis then seeks to assess the generalizability and structural manifestations of these insights across a large dataset, moving from micro-level understanding to macro/meso-level patterns. The objective, therefore, explicitly shifts from *verstehen* to identifying and analyzing recurrent structural patterns within the network as revealed by the ICIJ Offshore Leaks Database *and assuming we can generalise these structures* (more on that later under the methods section).

2.2 Conceptual foundations

This section outlines the necessary conceptual foundations that precede the concrete propositions asserted later in the thesis. These concepts presented here as being analytically requisite for the propositions developed in the subsequent section (2.2).

2.2.1 Global Wealth Chains and the Role of Intermediaries

To understand the significance of the intermediaries central to this thesis – the professional advisors, lawyers, accountants, and wealth managers operating within the offshore financial system – it is helpful to adopt an analytical framework that explicitly centers their role. The overall motivation for focusing on these actors stems from the "Global Wealth Chains" (GWC) approach.

As articulated by Seabrooke & Wigan (2014), this approach offers a distinct perspective compared to analyses focused on global value chains. They argue that: "While actors

in value chains share an interest in transparency and coordination, those in wealth chains thrive on rendering movements through the chain opaque. Wealth chains hide, obscure and relocate wealth to the extent that they break loose from the location of value creation and heighten inequality." Adopting this GWC lens necessitates an explicit focus on the intermediaries and professionals. These are the actors who develop and deploy the sophisticated financial and legal innovations required to sustain and manage the complex structures used to hold individual wealth offshore, often obscuring its origins and ownership.

Further elaborating on the socio-legal dynamics underpinning these chains, Seabrooke & Wigan (2022) emphasize the significance of socially constructed legal meaning. They write: "What is significant here is accepted legal assertions,. This happens within interpretative communities, where agreements on legal affordances are secured." The intermediaries operate within these communities, shaping and interpreting the boundaries of legal possibility. Seabrooke & Wigan (2022) also connect this to broader social valuations, noting that "An important element is that within such communities wealth confers honor, where the accrual and transfer of wealth without productive effort is held in high esteem (Veblen, 1899)."

Borrowing from the typology proposed in Seabrooke & Wigan (2022), the networks involving the intermediaries examined in this thesis align closely with their definition of "relational wealth chains." These are characterized as follows: "Relational wealth chains involve the exchange of complex tacit information, requiring high levels of explicit coordination. Strong trust relationships managed by prestige and status interactions make switching costs high." This description of relational wealth chains, emphasizing tacit knowledge, trust, coordination, and high switching costs due to the personal nature of the relationships, is highly with the ethnographic work of Harrington (2016) and how she outlines the structure and dynamics of the networks between wealth managers and their elite clients. This connection is also drawn by Seabrooke & Wigan (2022) themselves, who cite Harrington (2015) alongside related work by Beaverstock & Hall (2016) and de Carvalho & Seabrooke (2016) as evidence supporting the characteristics of relational wealth chains.

Furthermore, a developing body of literature situated within this GWC approach is examining how these professionals actively shape and navigate existing regulatory landscapes (e.g., Christen, 2021; Christensen & Seabrooke). This underscores the analytical purchase of the GWC framework for understanding the pivotal role of intermediaries not just as passive facilitators, but as active agents within the offshore system.

2.2.2 Weaponised Interdependence

The goal here is to outline the theoretical basis for viewing intermediaries not just as facilitators, but as potential points of leverage or vulnerability within the offshore system, thereby informing regulatory strategies.

A lens for such an analysis is provided by the concept of "weaponised interdependence," as developed by Farrell & Newman (2019). Their core argument posits that globalization, far from simply flattening the world or diminishing state power, has often created highly specific network topographies. These global networks—whether in finance, technology, or supply chains—are frequently characterized by asymmetric structures. Power, in this view, does not dissipate but rather concentrates at key hubs or 'chokepoints' within these networks. States or actors who control these chokepoints gain significant leverage over others who depend on access to the network, potentially allowing them to 'weaponize' this interdependence for strategic gain.

This logic of weaponised interdependence has been applied directly to the domain of global tax policy by Christensen (2024). He argues that states have often failed to fully harness the potential regulatory power they could wield by strategically targeting chokepoints within the networks facilitating tax avoidance and evasion. Among the key institutions Christensen (2024) identifies as potential chokepoints relevant to global tax policy are precisely the expert intermediaries – the lawyers, accountants, wealth managers, and corporate service providers – who are central to this thesis. Their specialized knowledge and gatekeeping function position them as critical nodes whose disruption could have widespread effects.

This perspective aligns with and provides a theoretical underpinning for findings across various studies highlighting the importance and potential vulnerability of the intermediary supply-side. Research emphasizing the role of intermediaries (e.g., Harrington 2016; Alstadsæter et al. 2019) implicitly points to their structural significance. For instance, Harrington's (2016) focus on trust-based relationships suggests that disrupting these specific intermediary nodes can create significant friction. Alstadsæter et al.'s (2019) supply-side explanation for high-end evasion similarly underscores the crucial role of these facilitators. More explicitly, recent work analyzing the network structures revealed by leaks, such as Chang et al. (2023), demonstrates the analytical purchase of focusing on these networks. While their specific study examined network structures to understand the effectiveness of sanction regimes against oligarchs, the underlying approach – analyzing network vulnerabilities by focusing on intermediary connections – is directly applicable to the broader question of regulating the offshore system for tax purposes.

All in all, understanding the network structure, particularly the role of intermediaries as potential chokepoints, reinforces the idea that the current state of offshore finance and associated tax evasion is, as Saez & Zucman (2019) argue in a related context, a

continued choice shaped by policy and enforcement priorities, rather than an immutable fact of nature.

2.2.3 Network Theory as a Lens for Understanding Illicit networks

To further contextualize the approach taken in this thesis, it is useful to briefly elaborate on how network studies have previously been employed to explore the structure and dynamics of analogous social and economic systems. The application of network analysis provides powerful tools for understanding complex relational patterns, information flows, and vulnerabilities within various types of networks, including those operating in clandestine or illicit domains.

The foundational work in social network analysis, such as Granovetter's (1973) seminal paper on the "strength of weak ties," laid the groundwork for understanding how network structures facilitate crucial processes like information diffusion and resource access. While initially focused on phenomena like job searching, these core insights into how different types of ties (strong vs. weak) and different network positions (e.g., bridges) shape outcomes have proven broadly applicable. Understanding the topology of connections is essential for identifying critical links, potential weaknesses, and influential actors within any network system. This foundational understanding extends to the analysis of illicit networks, where mapping relationships can reveal operational structures and vulnerabilities.

One of the prominent examples demonstrating the application of network analysis to understand illicit operations is the work of Morselli (2009). By examining specific cases, such as the CAVIAR network involved in cross-border drug smuggling, Morselli illustrates how network science concepts (like centrality measures, brokerage roles, and structural holes) can be used to dissect the organizational structure of criminal enterprises. Such analyses move beyond individual actors to understand the relational patterns that enable the illicit activity, potentially identifying key players or structural weaknesses that could be targeted for disruption.

More directly relevant to the subject matter and data source of this thesis, recent studies have begun applying network analysis to the large-scale datasets released by the ICIJ. Chang et al. (2023), for instance, utilized network methodologies on ICIJ data to specifically examine the effectiveness of sanction regimes against oligarchs, analyzing how their embeddedness within offshore networks influenced outcomes. Similarly, related work by the same authors ("Complex Systems of Secrecy," Chang et al. 2023) employed network perspectives to explore patterns related to the types of offshore instruments demanded by elites, linking structural features to strategic choices. These studies exemplify how network analysis can yield substantive insights from the complex relational data contained

within the ICIJ leaks, demonstrating its utility for exploring the offshore financial system.

The general principles and analytical techniques drawn upon in such studies are well-established within the broader field of network science, with standard references like Newman's (2010/2018) textbook providing comprehensive overviews of the underlying theory and methodologies. While this thesis may focus more on synthesis and proposition-building informed by network concepts rather than complex quantitative modeling, drawing upon this established body of work provides a robust conceptual and methodological grounding for analyzing the structure and significance of intermediary networks in offshore finance.

2.2.4 A Typology of Intermediaries and Their Role

To proceed with an analysis centered on the supply-side, it is essential to clarify conceptually what exactly is meant by an "intermediary" within the context of offshore finance. These actors play diverse roles in facilitating the creation, maintenance, and utilization of offshore structures. While specific studies, such as Harrington (2016), provide deep insights into the practices of particular intermediary types like wealth managers, a broader classification is useful for systematic analysis.

These are all what Hoang (2022) would call the "small spiders", the "High net worth indivduals" rather than the "Ultra-High net worth individiduals" sitting at the top of the food chain. Anything uncovered, in this respect is extremely limited, because they are able to further obfuscate their position.

For this purpose, this thesis builds upon the typology developed in a 2017 EU report examining the role of advisors and intermediaries as revealed in the Panama Papers. This framework, grounded in empirical observation of a major leak, categorizes intermediaries based on their primary area of expertise and function within the offshore ecosystem. Adopting this typology serves a dual purpose: it provides conceptual clarity for the subsequent discussion and offers a practical schema for efforts to classify the varied intermediary actors identified within the ICIJ dataset, thereby enriching the data for structural analysis.

Based on the EU (2017) framework, we can distinguish the following core types of intermediaries:

- **Tax Experts:** These intermediaries focus primarily on the tax implications of offshore structures. Their core function involves advising clients on tax planning strategies to minimize liabilities (potentially crossing into evasion) and ensuring compliance through the preparation of necessary tax documentation across relevant jurisdictions. This group can include accountants, auditors, and specialized tax advisors, whose advice may vary in aggressiveness.

- **Legal Experts:** This category encompasses professionals providing expertise on the legal design, establishment, and enforcement of offshore structures. Key activities include structuring entities to navigate or exploit laws in multiple jurisdictions, handling incorporation (often via licensed entities), drafting legal documents, arranging nominee services, and providing formal legal opinions or representation. This group includes regulated lawyers, who often have exclusive rights for certain actions like court representation, and potentially notaries involved in document formalization.
- **Administrators:** The primary role of administrators is the ongoing operational maintenance and financial record-keeping of offshore entities. This includes preparing financial accounts, potentially handling tax returns (overlapping with Tax Experts), managing day-to-day administrative tasks, and sometimes auditing accounts (though auditors require independence). Accountants often fall into this category, focusing on financial recording and reporting.
- **Investment Advisors:** Distinct from those setting up the structure, investment advisors focus on managing the assets held within the offshore entity. Their core function is to develop strategies for wealth preservation or growth using the financial instruments (or other assets like property, art, etc.) owned by the offshore structure. Their role is centered on asset management rather than the legal or tax architecture itself.

This typology provides a decent conceptual grounding for analyzing the distinct roles and potential influence of different supply-side actors within the offshore financial network.

2.2.5 Secrecy Strategies: Financial Instruments and Legal Innovations

Goal: Understanding the different financial instruments they use and how they can be innovated on, and used for different purposes. (Mainly Lafitte, 2024; Chang et al. 2023)

Most important type, Bearer instruments:

Harrington (2016) writes of Bearer instruments as follows: *In addition, a few offshore jurisdictions allow the use of "bearer shares," which are a way of issuing corporate stock without specifying a particular owner. Rather, the owner of a bearer share is literally whoever happens to be holding the stock certificate at any moment in time. This provides strong privacy protections, because as long as one does not have the shares in hand, one can say truthfully under oath, "I do not own that firm." And if any officers of the firm are ever questioned about its ownership, they can also truthfully say, "I don't know who owns the company, because bearer shares were issued." In other words, bearer shares make it impossible to know who owns a company, and that makes it impossible to assign legal responsibility for any taxes, fines, or debts the company incurs.*

2.3 Propositions

PLACEHOLDER

Data and Methodology

This section details the data sources and methodological approaches employed in this thesis. It begins by describing the primary data source, the ICIJ Offshore Leaks Database, outlining its structure, content, strengths, and limitations. Subsequently, it discusses the external datasets used to provide contextual information. Finally, it introduces a novel methodology utilizing agentic AI to enrich the classification of intermediaries within the icij dataset.

3.1 The ICIJ Offshore Leaks Database

The empirical core of this thesis rests upon the International Consortium of Investigative Journalists (ICIJ) Offshore Leaks Database. This resource serves as the primary data source, acting as a valuable, albeit imperfect, "proxy" for the opaque universe of offshore finance (cf. EU, 2017). The general idea underpinning its use here is that while any direct numerical estimates derived solely from the leaks (e.g., total wealth hidden) will surely be biased due to the data's inherent incompleteness, the qualitative nature of the interactions captured within the data – the patterns of relationships between clients, intermediaries, and jurisdictions – appears more reliable for understanding the structure and dynamics of the offshore system.

The use of the ICIJ Offshore Leaks Database for this type of research is increasingly established. For example, Alstadsæter et al. (2019), Londoño-Vélez & Ávila-Mahecha (2021), and Chang et al. (2023a; 2023b).

Some direct network analysis, but not much. Most relevant for our purposes is the work of Chang et al. (2023a; 2023b), as well as a more direct network study (Kejriwal & Dang, 2020) looking at the usual properties networks exhibit. They note:

"It was really unusual. The degree of fragmentation is something I have never seen before," said Kejriwal. "I'm not aware of any other network that has this kind of fragmentation."

The core datasets loaded for this analysis include:

- `nodes-entities.csv`: Information on offshore companies, trusts, foundations.

- `nodes-officers.csv`: Details on individuals or companies acting in official capacities (directors, shareholders, beneficiaries).
- `nodes-intermediaries.csv`: Data on firms or individuals facilitating the creation and management of offshore structures.
- `nodes-addresses.csv`: Physical address information linked to other nodes.
- `nodes-others.csv`: Nodes not fitting the primary categories.
- `relationships.csv`: The edge list defining connections between nodes, including the type of relationship.

A multi-modal and multi-relational graph, that often involved dealing with it at various level of abstractions, and breaking the dimensionality of it. For example, most network algorithms cannot deal with multi-relational graphs (in fact, a tetrapartite graph composed of addresses, entities, officers and intermediaries!), built and formulated as traditional adjacency matrices operations, so often switching between different representations of the data, ranging from granular edge lists with the full ontology of the data model, to those squashed down into a format that can be represented as a single adjacency matrix (i.e. an edge list with only a single type of source- and end node).

Usual trick for dealing with bipartite graphs is projecting it down by connecting two nodes if they share a common node type, that one wants to eliminate. For example, for getting rid of address node type, relations like the following would be transformed as such:

Intermediary1 – Address – Intermediary2 -> Intermediary1 – Intermediary2

This process can be done iteratively on all node types that one wants to eliminate: Intermediary1 – Entity – Address – Officer – Intermediary2 -> Intermediary1 – Intermediary2

And from here, traditional network algorithms calculating the likes of eigenvector centrality etc. can be applied.

Whenever moving between the different representations of the graph, will be made clear in the empirical section.

Note: ICIJ data is in principle directed, but in practice this is not especially important. Here, treated solely as an undirected graph.

3.2 External Data Sources

To contextualize the patterns observed within the ICIJ data, several external data sources are employed.

A key resource is the Historical Tax Havens Database (HTHD) developed by Laffitte (2024). This dataset documents the historical evolution of "offshore legal architecture," tracking the adoption of specific legal technologies (e.g., banking secrecy, IBCs) across tax havens over time. This dataset will be utilized to explore whether specific patterns observed in the ICIJ data – such as the prevalence of certain offshore instruments or shifts in intermediary activity – align temporally with the historical innovations documented in the HTHD.

The World Justice Project (WJP) Rule of Law Index provides comprehensive country-level metrics on governance. Its specific use is to investigate potential correlations between the home country's rule of law environment and the patterns of specialization or network positioning observed among the intermediaries serving clients from that country.

VDEM (Varieties of Democracy) Regime Type Data will be used exactly analogously.

Data from the World Inequality Database (WID), specifically metrics on wealth inequality at the country level, will also be incorporated. This serves primarily to see if we can confirm some of the comparative statics Alstadsæter and Zucman derive, trying to verify whether there's anything to their supply-side model.

3.3 Using Agentic AI to Scrape Data on Intermediaries.

A significant challenge in utilizing the ICIJ data for the purposes of this thesis is that intermediaries are often classified generically within the database. To analyze the specific roles and potential influence of different types of intermediaries, as outlined in the typology adapted from the EU (2017) paper (see Section 2.1.4), a more granular classification is required. To achieve this classification at scale, an approach employing agentic AI is utilized.

The core idea is to use an AI agent loop to automate the process of gathering information about and classifying the intermediaries listed in the ICIJ data. The basic workflow is illustrated in Figure 3.1.

In brief, the process involves an AI agent orchestrating online searches for each intermediary identified in the ICIJ data. It begins with generic searches, reads and interprets the initial results, and then formulates more specific search queries based on the information discovered or identified as lacking. This iterative process involves up to three search queries per intermediary, scouring the top 15 most relevant web results identified through query-result embedding similarity using the Tavily Search API (though the tool is relatively generic and its specific choice is not critical to the methodology). This effectively replaces the time-consuming need for manual searching of the intermediaries.

Based on the information gathered, the AI agent then classifies the intermediary ac-

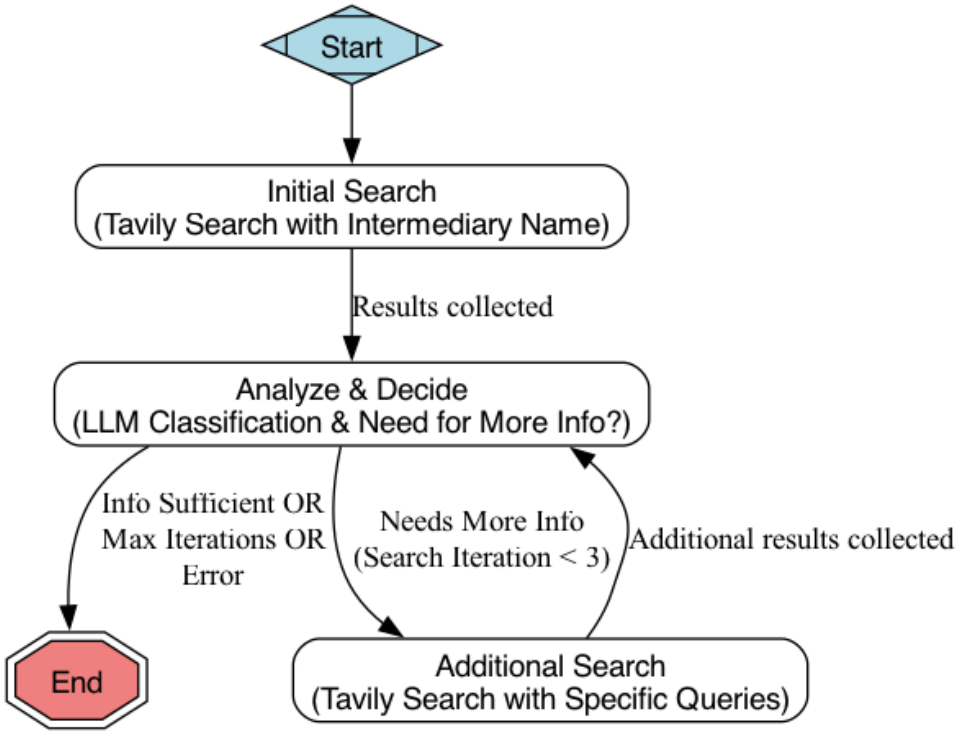

Figure 3.1: Agent Setup for Intermediary Classification

cording to the EU (2017) typology (Tax Expert, Legal Expert, Administrator, Investment Advisor), adding a few additional relevant fields (e.g., specific job title). Crucially, the agent also provides a confidence score for its classification judgment, allowing for filtering or weighting in subsequent analyses.

There are a few obvious limitations associated with this approach that warrant discussion:

Temporal Misalignment: A primary concern is that all online searches are conducted based on information available today, whereas the ICIJ data pertains to activities that may have occurred years or even decades prior. This introduces two potential issues:

1. The process might be biased towards identifying intermediaries who are still active or have a significant online presence currently.

2. It implicitly assumes that the role an intermediary plays today (as reflected online) is equivalent to the role they played at the time relevant to the ICIJ data.

Addressing the Limitations: While these issues are real, they are not prohibitive for their use in this thesis.

1. Regarding the first point (bias towards current intermediaries), this primarily impacts the coverage or statistical power of the classification – we may only be able to confidently classify a subset (e.g., 50%) of all intermediaries. This is acceptable, provided the unclassifiable intermediaries are not systematically different in ways

that correlate with the research questions. The issue becomes problematic only if there is a systematic bias in identifiability across types (e.g., if it is inherently much harder to find information online about legal experts compared to tax experts due to differing needs for discretion or public visibility). The significant threat here is the bias in who provides public information - it will only be the intermediaries whose activities are not inherently illegal. A bias towards, in a sense, the least dangerous intermediaries as those being revealed.

2. Regarding the second point (role stability), the assumption that roles remain consistent is arguably less problematic. Given the highly specialized nature of functions like tax advisory, legal structuring, administration, and investment management within the offshore context, and the considerable barriers to entry (qualifications, reputation, networks) for each, frequent switching between these core roles by individuals or firms seems relatively unlikely.

In my view, it is the only pragmatically feasible method to do this given the constraints of this thesis.

To instruct the AI agent on how to perform the classification and the specific structure of the information to return, the following prompt template is utilized. This prompt defines the categories, provides keywords for guidance, and specifies the desired output fields. The agent's output for each intermediary is a structured data record, typically resembling a JSON object or a Python dictionary, which includes the fields detailed in the prompt.

Classification Prompt

The core prompt provided to the AI agent for classification is as follows (where `{intermediary_name}` and `{log_summary_for_classification}` are dynamically inserted):

```
Classify the intermediary: {intermediary_name}

Based *only* on the information gathered in the following search log.
{log_summary_for_classification}

Classify this intermediary into ONE of these categories based on their
likely primary role in offshore activities:
- Tax Expert: Focuses on tax planning, compliance, advisory. Keywords:
  tax advisory, international tax, tax compliance, tax returns,
  transfer pricing, VAT, tax structuring.
- Legal Expert: Focuses on legal structuring, compliance, incorporation,
  representation. Keywords: legal services, corporate law, entity formation,
```

```
  incorporation, contracts, litigation, legal opinions, regulatory
  compliance, M&A legal, lawyer, attorney, solicitor.
- Administrator: Focuses on accounting, auditing, financial reporting,
  company administration. Keywords: accounting, bookkeeping, audit,
  financial statements, reporting, company secretarial, payroll,
  administration services, domiciliation, accountant, auditor.
- Investment Advisor: Focuses on managing financial assets and investments.
  Keywords: investment management, wealth management, asset management,
  portfolio management, financial planning, investment strategy,
  securities, funds, financial advisor.

Provide a structured classification including:
- classification (Enum: Tax Expert, Legal Expert, Administrator, Investment Advisor)
- role_muddled (bool: true if the role seems mixed or unclear)
- role_muddled_reasoning (str: explanation if role_muddled is true)
- is_individual (bool: based on the name and findings, is this likely a person?)
- job_title (str: inferred job title if possible, e.g., "Lawyer", "Accountant",
  "Director", or "Unknown")
- confidence (Enum: Low, High - Use Low if evidence is sparse, contradictory,
  or confidence in the source/relevance is low)
- justification (str: detailed reasoning for the classification, referencing the
  search log)
- key_evidence (list[str]: specific snippets or findings from the search
  results supporting the classification)

Analyze the content of the search results carefully. Prioritize information
directly describing the intermediary's services or professional role.
```

Examples of Dynamic Search and Structured Output

The agent's search process is dynamic. It begins with a general query (the intermediary's name) and, based on the retrieved information's relevance and completeness, may formulate up to two additional, more specific queries. For instance, if initial results for a company are vague, subsequent queries might include terms like "services offered" or "business activity." The classification is then made based on the entirety of the gathered search logs.

Sometimes it's just not possible to find anything useful, hence the confidence section. Any cases where 'confidence' is low, they are excluded from the analysis sections.

The output for each intermediary is a structured record. While the 'key_evidence'

field in the prompt requests specific snippets, for comprehensiveness in these examples, it contains the full, somewhat verbose, search log detailing each iteration of the dynamic search process.

The following examples illustrate this process and the resulting structured output, taken directly from the enrichment process. The first provides an example of where it's impossible to find clear information linked to the interemediary resulting in low confidence, while the second is a clear case of a legal expert.

Example 1: OPULENT OPAL HOLDINGS LIMITED

Dynamic Search Summary:

- **Iteration 0 Query:** 'OPULENT OPAL HOLDINGS LIMITED'
- **Findings:** Results were ambiguous, pointing to "OPULENT OPAL CONSTRUCTION LIMITED," a subsidiary link to "Opal Holdings Pte Ltd" (itself a subsidiary of CapitaLand, suggesting a holding structure), and unrelated products (paint, home decor).
- **Iteration 1 Query:** 'OPULENT OPAL HOLDINGS LIMITED primary business activities'
- **Findings:** Retrieved general information on Bank Secrecy Act compliance for businesses, potentially triggered by "holdings," and another mention of the CapitaLand subsidiary link.
- **Iteration 2 Query:** 'OPULENT OPAL HOLDINGS LIMITED services offered'
- **Findings:** Largely irrelevant results (gemstones, furniture), again highlighting the difficulty in pinpointing specific services for this entity.

Structured Output (Simplified Representation):

- **intermediary_name:** OPULENT OPAL HOLDINGS LIMITED
- **classification:** Administrator
- **confidence:** Low
- **role_muddled:** True
- **role_muddled_reasoning:** "The role is muddled as the search results are ambiguous and could indicate construction, holding company activities, or other possibilities. There is no clear indication of their primary business activities."
- **is_individual:** False

- **job_title:** Unknown
- **justification:** "The search results provide limited information about OPULENT OPAL HOLDINGS LIMITED. One result mentions 'OPULENT OPAL CONSTRUCTION LIMITED' which suggests involvement in construction, potentially as a specialty trade contractor. Another result indicates that 'Opal Holdings Pte Ltd' is a subsidiary of CapitaLand, implying a holding company structure. The lack of direct evidence about their services or role leads to a low confidence classification as an administrator, potentially involved in company administration or holding activities."
- **key_evidence (Full Search Log):** (Content of the JSON-like string representing the three search iterations and their results, as provided in the original example, would be detailed here. For brevity, it is summarized above and omitted in this representation.)

Example 2: Cheo Yeoh & Associates

Dynamic Search Summary:

- **Iteration 0 Query:** ‘Cheo Yeoh & Associates‘
- **Findings:** Immediately yielded relevant results identifying the entity as "CHEO YEOH & ASSOCIATES LLC," a Singapore-based company providing legal services, with details on its registration and directors.
- **Iteration 1 Query:** ‘Cheo Yeoh & Associates services‘
- **Findings:** Reinforced the legal nature, with results like "Cheo Yeoh & Associates provides legal services" and official listings specifying "LEGAL ACTIVITIES."
- **Iteration 2 Query:** ‘Cheo Yeoh & Associates client testimonials‘ (or similar refinement if needed)
- **Findings:** Could uncover legal case mentions, lawyer profiles, or further business directory listings, solidifying its role. (The provided log indicates successful identification primarily from the first two iterations).

Structured Output (Simplified Representation):

- **intermediary_name:** Cheo Yeoh & Associates
- **classification:** Legal Expert
- **confidence:** High

- **role_muddled:** False
- **role_muddled_reasoning:** (Empty or N/A)
- **is_individual:** False
- **job_title:** Lawyer (Implicitly, as a law firm)
- **justification:** "Cheo Yeoh & Associates LLC provides legal services in Singapore. The company's principal activity is LEGAL ACTIVITIES (EXCLUDING ONLINE MARKETPLACES). Johnny Cheo is a Director at the firm. The firm has been operating for 23 years. The search results consistently point to legal services."
- **key_evidence (Full Search Log):** (Content of the JSON-like string representing the search iterations and their results, as provided in the original example, would be detailed here. For brevity, it is summarized above and omitted in this representation.)

3.4 Use of LLMs

3.5 Use of LLMs in the Broader Paper

LLMs have also been used to polish the text of this thesis and used for idea generation. Used Google Gemini models mainly, with the seed configured to 42:

- gemini-2.5-pro-preview-05-06
- gemini-2.5-pro-experimental-03-25

Empirical Analysis

Generally, in the sources that have been leaked, a gradually increasing number of incorporations. Entity creation across different data leaks. Figure 4.1 illustrates the number of entity incorporations over time, segmented by the source investigation (e.g., Panama Papers, Pandora Papers).

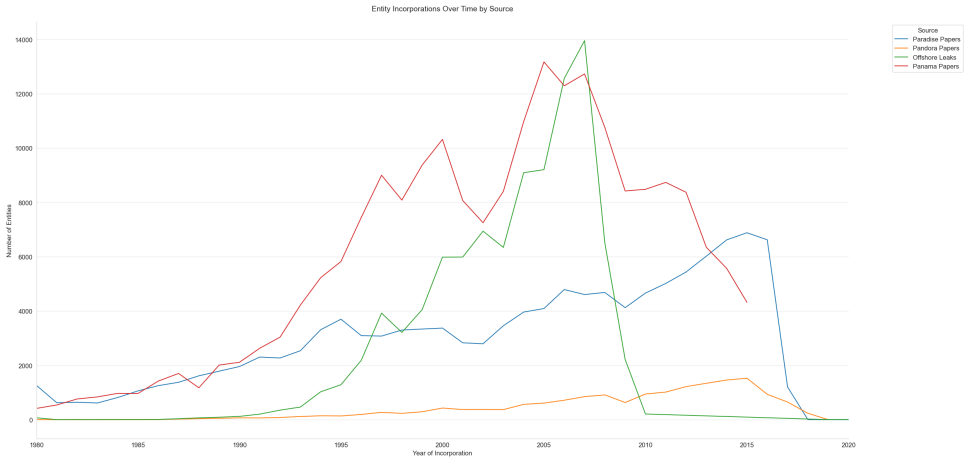

Figure 4.1: Entity Incorporations Over Time by Source Investigation.

Because of the power-law distribution intermediaries follow in terms of their connections (see Figure 4.2), we can explain a significant portion of the network activity by enriching the top fraction of intermediaries. Specifically, enriching the top approximately 750 intermediaries (ranked by the number of entities they are connected to) provides substantial coverage of 60% entities, as shown in Figure 4.3.

The geographic distribution and coverage of different node types (intermediaries, entities, officers) provide further context. Figures 4.4 and 4.5 show the primary locations of intermediaries and the coverage achieved by focusing on the top locations. Similarly, Figures 4.6 and 4.7 depict the distribution and coverage for entities, while Figures 4.8 and 4.9 do the same for officers.

The common thread across all of them, is that all are incredibly concentrated in a limited number of countries.

The enrichment process involved classifying intermediaries based on their likely role (e.g., Administrator, Legal Expert). Figure 4.10 compares the distribution of these clas-

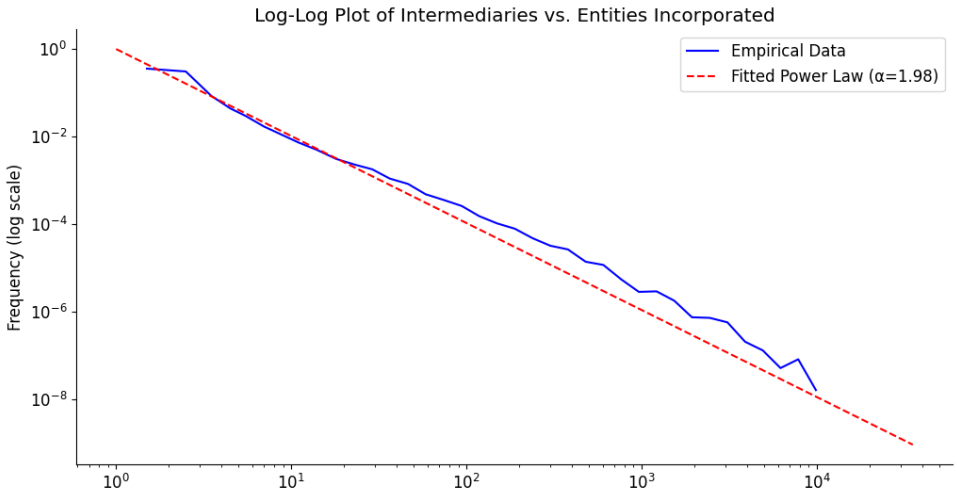

Figure 4.2: Illustration of the Power Law Distribution of Intermediary Connections (Degree).

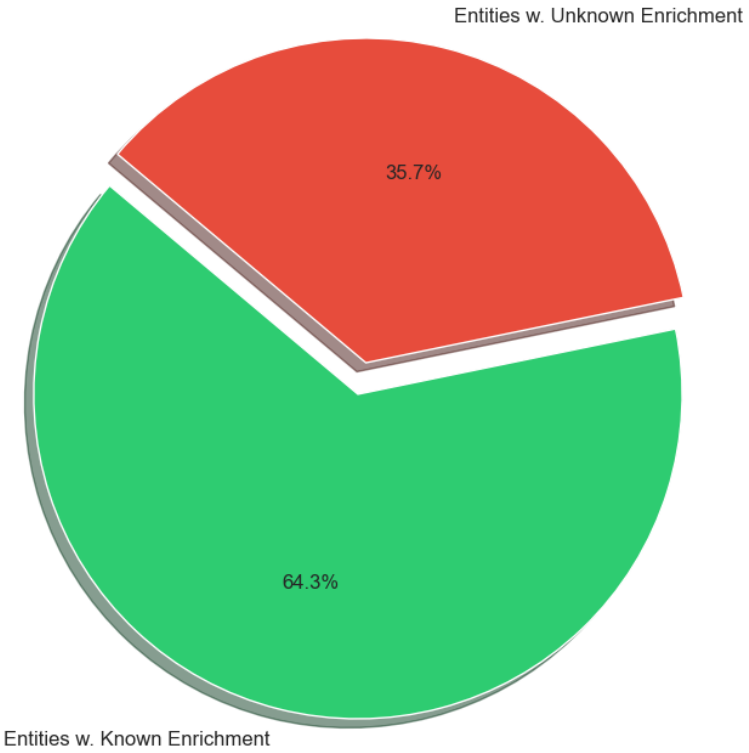

Figure 4.3: Coverage of Intermediary Connections Explained by the Enriched Top Intermediaries.

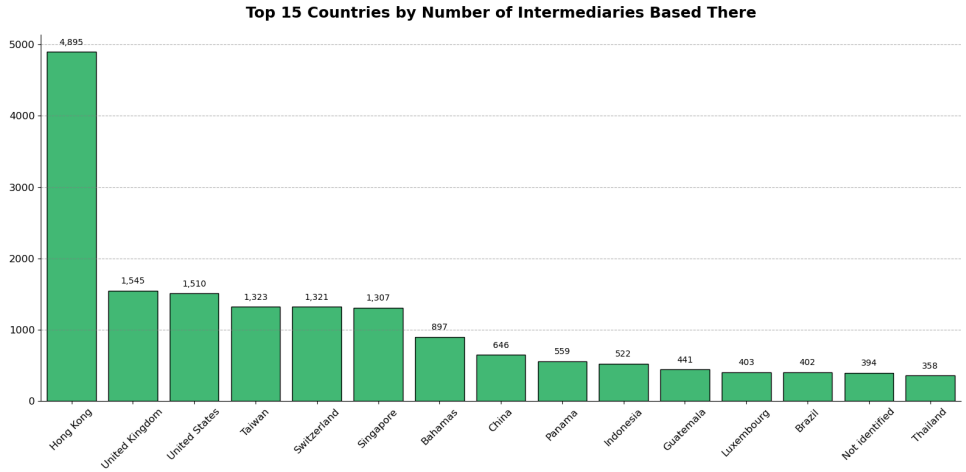

Figure 4.4: Geographic Distribution of Intermediaries (Top Locations).

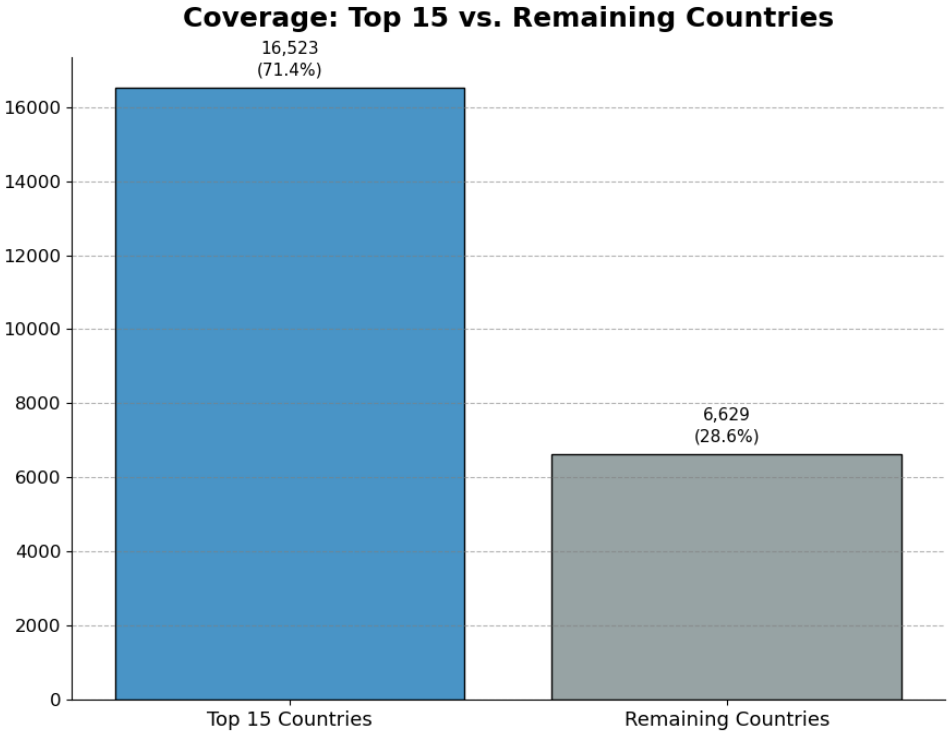

Figure 4.5: Coverage of Intermediaries by Top Geographic Locations.

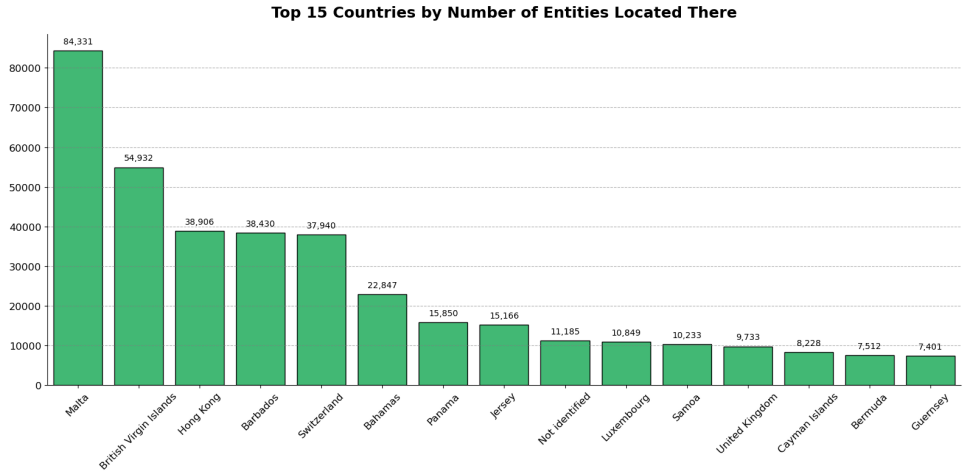

Figure 4.6: Geographic Distribution of Entities (Top Jurisdictions/Locations).

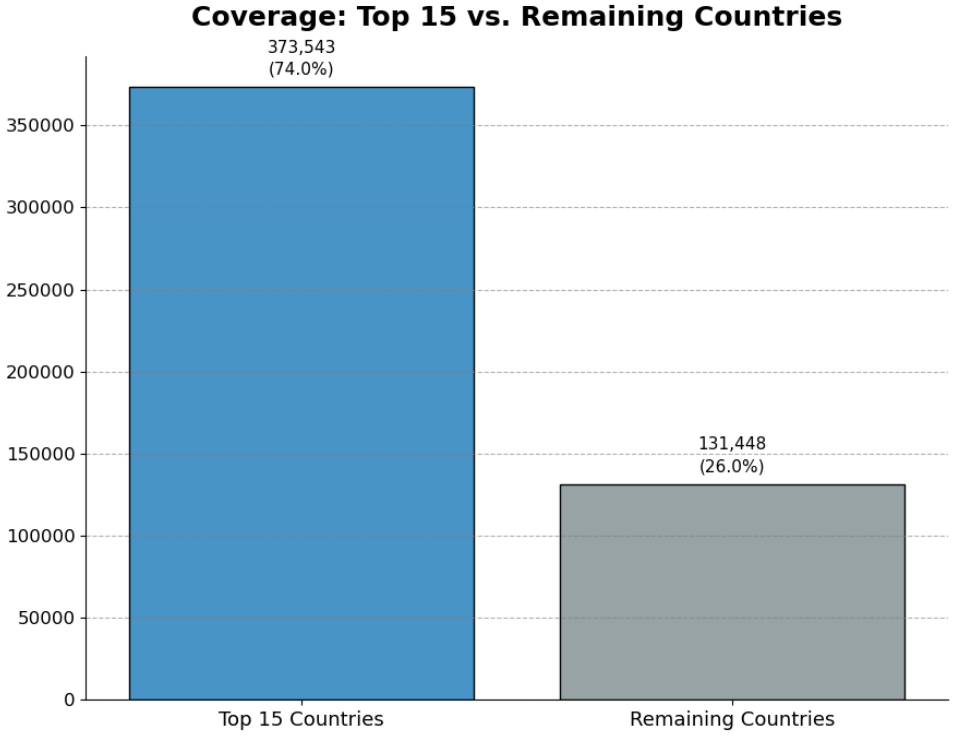

Figure 4.7: Coverage of Entities by Top Geographic Locations.

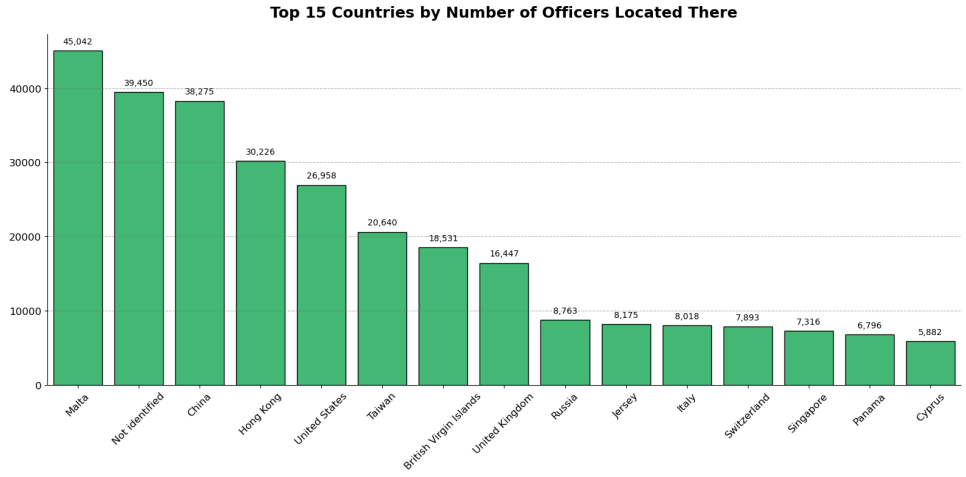

Figure 4.8: Geographic Distribution of Officers (Top Locations).

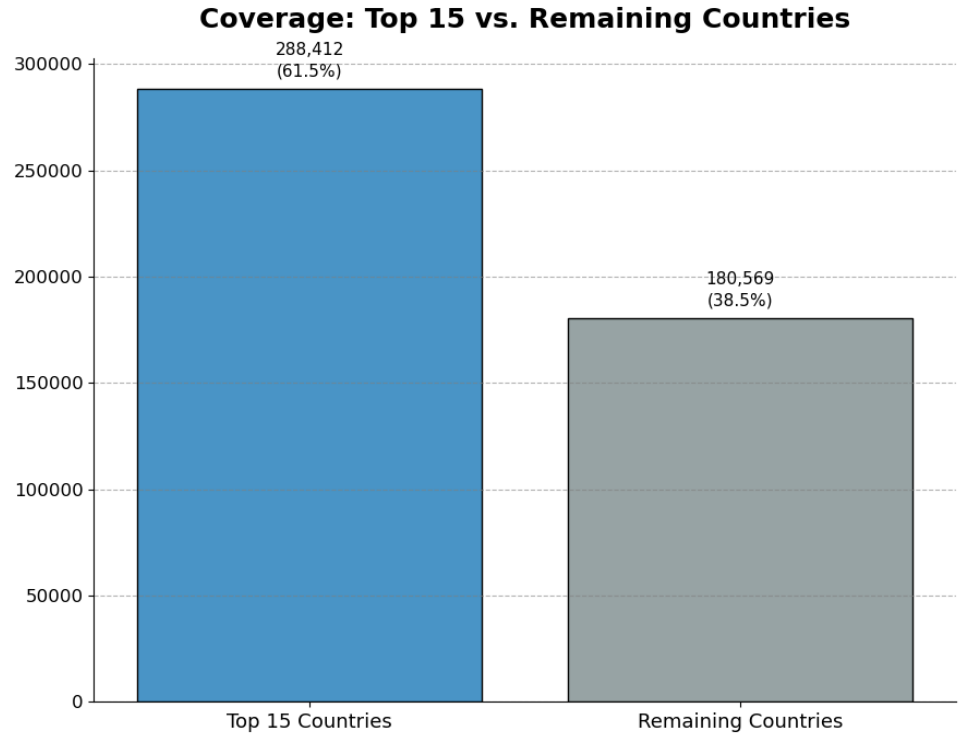

Figure 4.9: Coverage of Officers by Top Geographic Locations.

sifications within the random sample versus the sample of top 5% intermediaries.

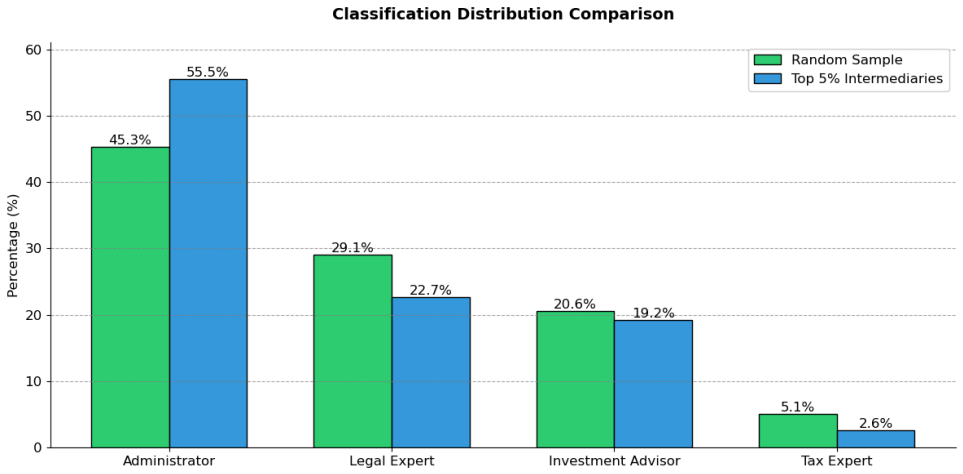

Figure 4.10: Comparison of Intermediary Classification Distributions (Random Sample vs. Top 5%).

We observe distinct degree distributions based on these classifications. Figure 4.11 shows the Cumulative Distribution Function (CDF) of intermediary degrees for each classification, while Figure 4.12 presents this comparison using boxplots.

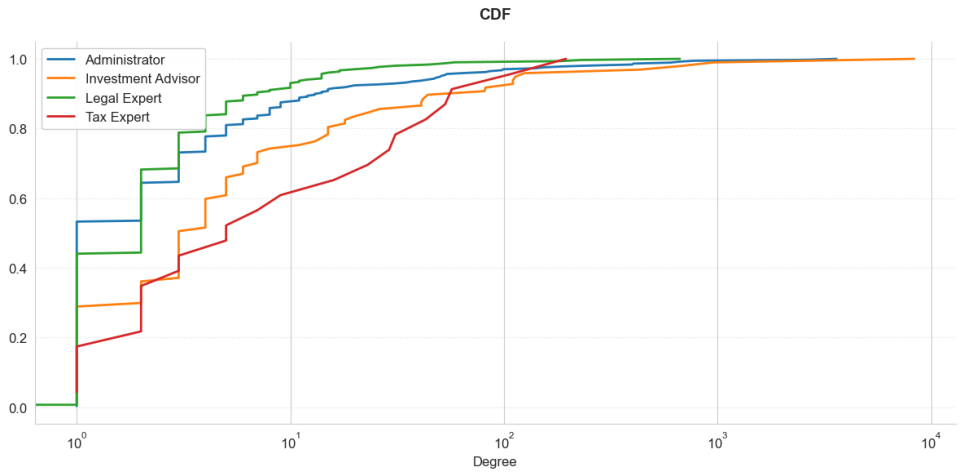

Figure 4.11: Cumulative Distribution Function (CDF) of Intermediary Degree by Classification.

These classification patterns also vary geographically, as illustrated in Figure 4.13, which likely shows the prevalence of different intermediary types across key countries.

Also, differing degree distributions, tested signifiance for.

Using a Negative Binomial GLM; cannot use Poisson because of overdispersion.

Basic setup:

Random Component: The response Y_i (degree of intermediary i) follows a Negative Binomial distribution, $Y_i \sim \text{NB}(\mu_i, \theta)$, accommodating count data with overdispersion where $\text{Var}(Y_i) = \mu_i + \mu_i^2/\theta$.

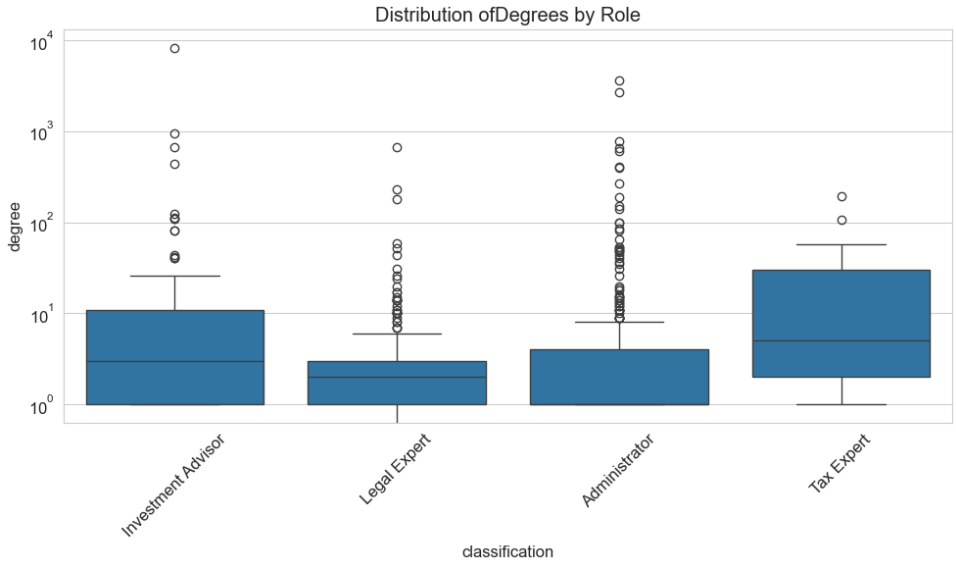

Figure 4.12: Boxplot of Intermediary Degree by Classification.

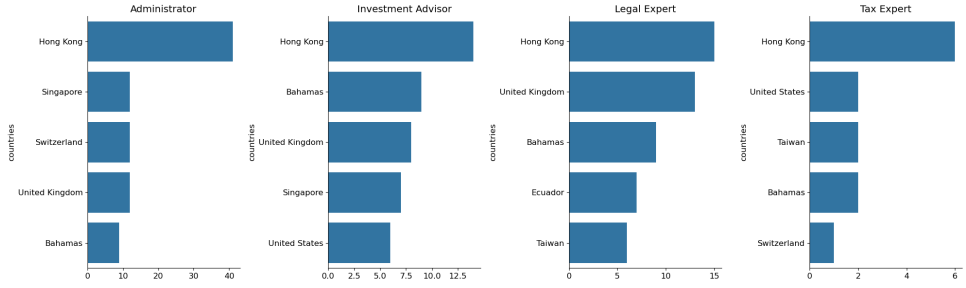

Figure 4.13: Distribution of Intermediary Classifications across Top Countries.

Link Function: The expected value $\mu_i = E[Y_i]$ is related to the linear predictor $\eta_i = \mathbf{x}_i^T \boldsymbol{\beta}$ via the **natural logarithm** link function:

$$\ln(\mu_i) = \eta_i \quad \Longrightarrow \quad \mu_i = \exp(\mathbf{x}_i^T \boldsymbol{\beta})$$

where $\mathbf{x}_i$ includes dummy variables for the intermediary classification (with 'Administrator' likely as the baseline) and $\boldsymbol{\beta}$ are the coefficients to be estimated.

Since there is no closed-form solution for the parameters $\boldsymbol{\beta}$ and the dispersion parameter θ, we numerically optimise the log-likelihood function:

$$\ell(\boldsymbol{\beta}, \theta) = \sum_{i=1}^{n} \ln \left[P(Y_i = y_i | \mu_i = \exp(\mathbf{x}_i^T \boldsymbol{\beta}), \theta) \right]$$

Results:

GLM Results (Negative Binomial):

Generalized Linear Model Regression Results

Dep. Variable:	degree	No. Observations:	790
Model:	GLM	Df Residuals:	786
Model Family:	NegativeBinomial	Df Model:	3
Link Function:	Log	Scale:	1.0000
Method:	IRLS	Log-Likelihood:	-3216.9
Date:	Mon, 12 May 2025	Deviance:	2927.8
Time:	13:00:11	Pearson chi2:	3.51e+04
No. Iterations:	12	Pseudo R-squ. (CS):	0.5825
Covariance Type:	nonrobust		

	coef	std err	z	P>|z|
Intercept	3.4632	0.053	65.419	0.000
C(classification)[T.Investment Advisor]	1.3162	0.115	11.457	0.000
C(classification)[T.Legal Expert]	-1.5206	0.081	-18.734	0.000
C(classification)[T.Tax Expert]	-0.2034	0.219	-0.929	0.353

The GLM results indicate significant differences in degree based on classification. Compared to the baseline (likely 'Administrator'), 'Investment Advisor' intermediaries have significantly higher degrees, while 'Legal Expert' intermediaries have significantly lower degrees. The 'Tax Expert' classification does not show a statistically significant difference from the baseline in this model. These findings are further corroborated by an Ordinary Least Squares (OLS) regression using the logarithm of the degree (log(degree)) as the target variable (results not shown here but confirm the direction and significance of the effects).

While static comparisons are (moderately) informative, the temporal progress of the network structure offers more dynamic insights. Figure 4.14 presents a case study focused on Cyprus, illustrating how the connections between different types of intermediaries and the locations (jurisdictions/countries) of the entities they service evolve over time.

Starting with all intermediaries from Cyprus, and then grouping entities together based on where they are 1) based in terms of business activity ('countries' field provided by ICIJ) and the jurisdiction whose rules they follow ('jurisdiction' field provided by ICIJ)

Two things to note:

1. Corroborates the well-documented picture of a gradually deepening network tied to Russia, as more and more entities connected to it, and

2. Primary intermediaries that are first needed are Administrators, and then more peripheral intermediary roles become involved (e.g., Legal Experts, Tax Experts). Too sparse to conclude it from this case study alone, but consistent with the idea that the network starts with a core of essential roles and then expands to include more specialized functions as the network grows.

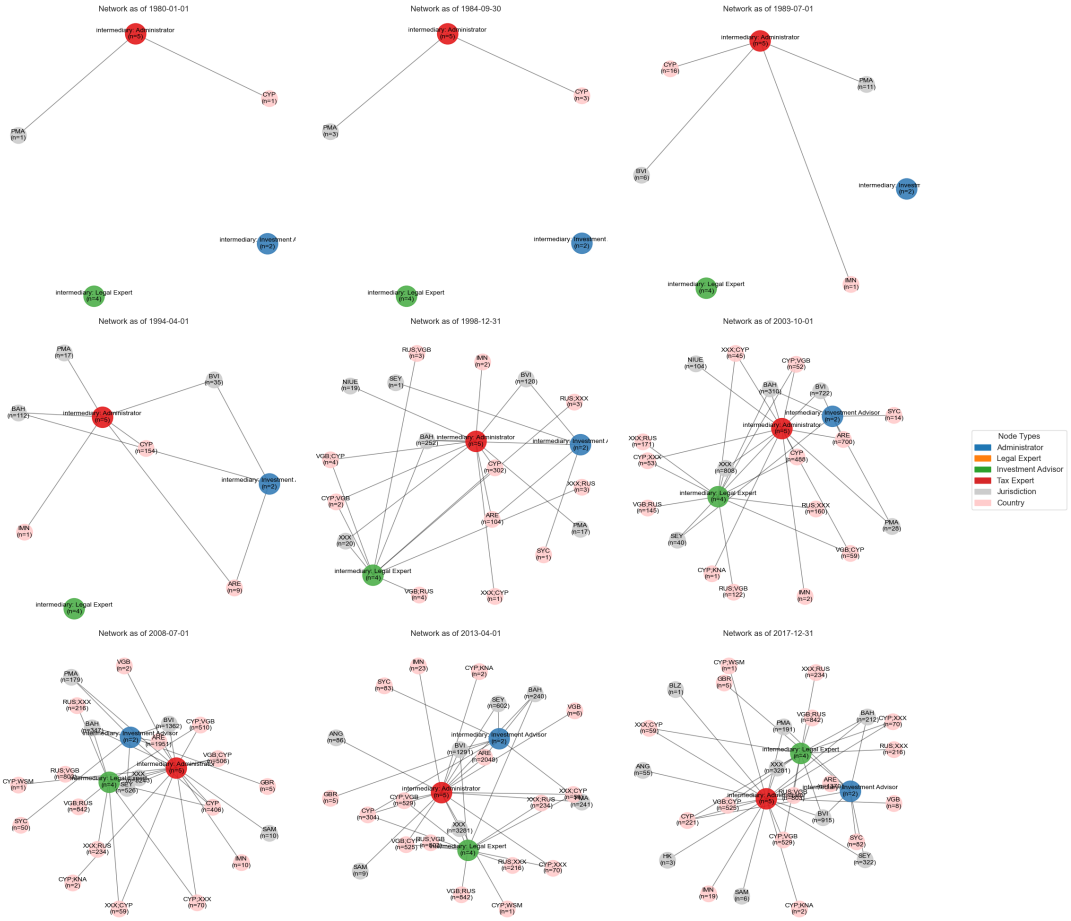

Figure 4.14: Temporal Evolution of the Intermediary-Entity Network (Cyprus Case Study). Shows connections between intermediary types (based in Cyprus) and the locations of entities they manage at different time points.

Discussion

Conclusion

Appendix