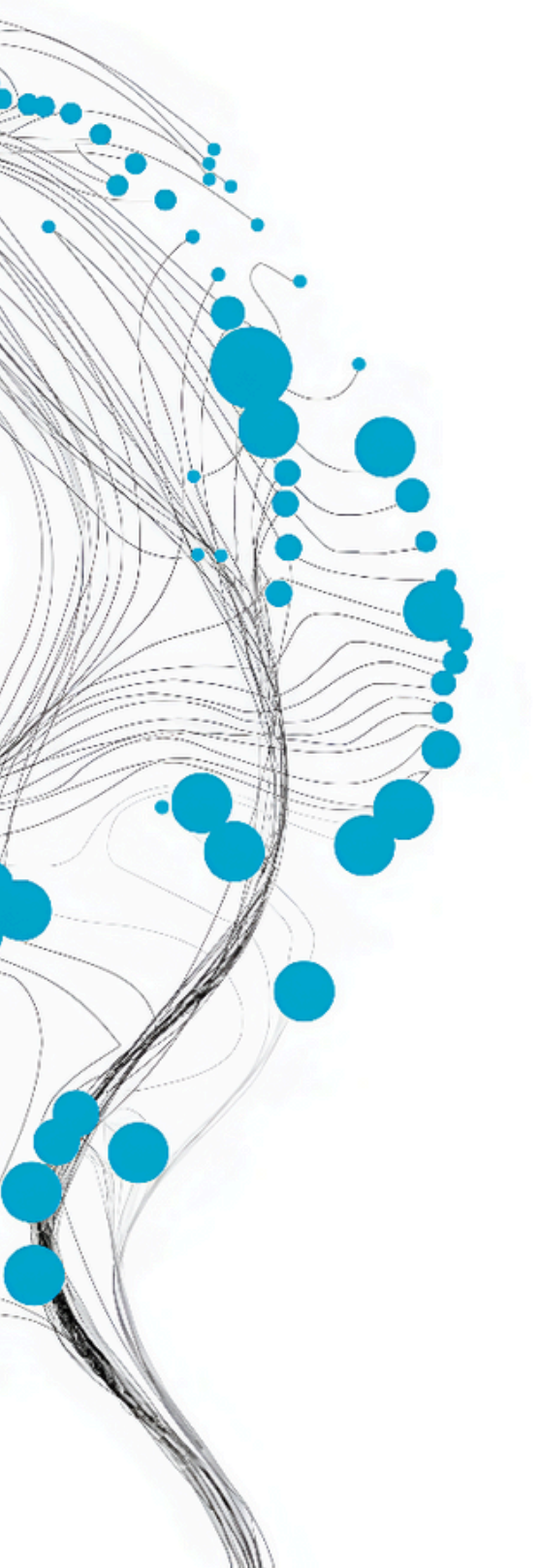


MASTER THESIS

**FAMILY OWNERSHIP AND FIRM
PERFORMANCE: A CROSS-REGIONAL
ANALYSIS OF THE USA AND EUROPE**



Author: Maurits Weenink
Study: Business Administration
Faculty: Behavioral, Management and Social
Sciences
Date: 23 July 2025

Graduation Committee Members:
Prof. dr. ir. A. Bruggink
H. Kroon.

UNIVERSITY OF TWENTE.

Abstract,

Family firms dominate both developed and emerging economies, accounting for a significant share of GDP and employment. Their unique ownership structure often fosters long-term commitment, relational capital, and intergenerational vision. However, the academic literature remains divided on whether family ownership enhances or hinders firm performance. On one hand, family involvement may reduce agency costs and foster stewardship, on the other, it may lead to nepotism, underinvestment in innovation, or misalignment with minority shareholders.

This complexity is further amplified by regional institutional contexts. In Europe, family firms are embedded in stakeholder-oriented governance systems that emphasize long-term relationships, employee involvement, and collective decision-making. In contrast, the United States features a more market-driven governance model characterized by dispersed ownership, shareholder primacy, and performance-based incentives. These contrasting environments could either amplify or constrain the benefits of family ownership, yet direct cross-regional comparisons remain scarce.

This thesis addresses that gap by asking: Are there significant differences in the relationship between performance of family versus non-family firms between the USA and Europe? To answer this, a structured analysis is conducted using a dataset of 4,026 publicly listed firms from Orbis, covering the period 2021–2024. Firm performance is measured using both Return on Assets (ROA) and Tobin's Q, capturing internal profitability and market valuation, respectively. To identify family firms, a multi-step classification method is employed where firms are first flagged if one or more named individuals or families hold a certain amount of direct or total ownership, additional classification is applied based on either a name match between shareholders and board members or the presence of family-related keywords in shareholder names. Performance differences are analyzed using ordinary least squares (OLS) regressions, incorporating robust standard errors and control variables for firm size, age, industry, growth, and leverage.

The findings show that family ownership is positively and significantly associated with accounting-based performance (ROA) in both regions, but the effect is notably stronger in the United States. No significant association is found between family ownership and Tobin's Q in either region, suggesting that the internal profitability advantages of family firms are not consistently priced into market valuations. These results challenge the widespread assumption that Europe's stakeholder-oriented institutions naturally enhance family firm performance. While European governance frameworks emphasize long-term stakeholder relationships, employee involvement, and social responsibility, these features may not always translate into enhanced financial outcomes. Instead, U.S. family firms appear more capable of leveraging their unique resources such as tacit knowledge, reputational capital, and long-term commitment into measurable profitability. This may be attributed to stronger competitive pressures, more transparent financial systems, clearer alignment between ownership and management, and a greater tendency to professionalize leadership structures. In the U.S. context, these factors seem to create an environment in which the advantages of family ownership are more effectively leveraged.

These findings imply that family firm performance is not solely shaped by ownership structure, but by how effectively that structure interacts with regional institutions. Policymakers and investors should therefore avoid assuming universal benefits or risks of family control, and family firm leaders should focus on converting internal strengths into externally visible performance through strategic governance and communication practices.

Keywords

Family Firms, Corporate Governance, Firm Performance, Tobin's Q, Return on Assets (ROA), Linear Regression, Europe, United States

Contents

1	Introduction	5
1.1	Importance	5
1.2	Problem Statement	5
1.3	Research Question	6
1.4	Contribution	6
2	Theoretical Framework	7
2.1	Definition Family Firms	7
2.2	Corporate Governance	7
2.2.1	Agency Theory	7
2.2.2	Stewardship Theory	8
2.2.3	Resource Based View	9
2.2.4	Institutional Theory	9
2.3	Regional Difference	10
2.4	Hypotheses	11
3	Methodology	11
3.1	Data Collection	11
3.2	Family Ownership Determination	11
3.3	Variables	12
3.3.1	Dependent Variables	12
3.3.2	Independent Variables	12
3.3.3	Control Variables	13
3.4	Sample	14
3.5	Regression Model	15
4	Results	16
4.1	Descriptive Analysis	16
4.1.1	Descriptive Statistics Sample	16
4.1.2	Descriptive Statistics of Family- and Non-Family Firms	16
4.1.3	Descriptive Statistics of Family- and Non-Family Firms in the EU	17
4.1.4	Descriptive Statistics of Family- and Non-Family Firms in the USA	18
4.2	Correlation Matrix	20
4.3	Linear Regression	20
4.3.1	Linear Regression	20
4.3.2	Linear Regression per Region	21
4.4	Potential Endogeneity	22
5	Conclusion	24
5.1	Hypothesis Testing	24
5.2	Interpretation	24
5.3	Implications	25
6	Discussion	26
6.1	Limitations	26
6.2	Future Research	26
7	References	27

8	Appendices	31
8.1	Appendix 1: Data Extraction	31
8.1.1	Appendix 1A: Variables	31
8.1.2	Appendix 1B: Method	32
8.1.3	Appendix 1C: Filters Applied	32
8.2	Appendix 2: Countries	33
8.3	Appendix 3: Industry Categories	34
8.4	Appendix 4: Financial Calculations	35
8.5	Appendix 5: Ruleset	36
8.6	Appendix 6: Frequencies	37
8.7	Appendix 7: Codes	37
8.7.1	Appendix 7A: Dataset Assembly	37
8.7.2	Appendix 7B: Datasample Assembly	40
8.7.3	Appendix 7C: Datasample Enhanced	42
8.7.4	Appendix 7D: Datasample Categorized	43
8.7.5	Appendix 7E: Descriptive Analysis	46
8.7.6	Appendix 7F: Linear Regression	48
8.7.7	Appendix 7G: Correlation Matrix	51

1 Introduction

1.1 Importance

According to Elstrodt and Pouillet (2014), family businesses, which account for 70–90% of gross domestic product (GDP) in some regions and contribute 54% to the Dutch GDP, are the most common form of business globally. In both developing and developed countries, family firms are a widespread phenomenon (Khanna and Rivkin, 2001). Family firms share significant and common business characteristics worldwide (Bunkanwanicha et al., 2013; Villalonga and Amit, 2006). However, whether family ownership enhances or diminishes firm performance remains a controversial topic. Chrisman et al., 2012 argue that family ties, through ownership and management, influence firms by prioritizing family goals and agendas. This approach may lead to the allocation of firm resources toward personal family objectives, potentially impacting shareholder wealth (Carney, 2005; E.-T. Chen et al., 2013; Chrisman et al., 2013). Such dynamics often create conflicts between family and non-family shareholders (Madison et al., 2016). Moreover, family-controlled firms may adopt policies that favor their own interests, at the expense of minority shareholders (Yeh and Woidtke, 2005). For instance, studies by Y. Liu (2011), Q. Liu et al. (2015) and Ozkan and Ozkan (2004) highlight that family firms tend to hold higher levels of cash, which can harm firm value. Similarly, Kuan et al. (2011) suggest that corporate governance practices influence cash holdings in both family and non-family-owned firms. Prior research indicates that family firms are often better positioned to seize business opportunities, align managerial and shareholder interests, and obtain financing under more favorable conditions (Hegde et al., 2020). These advantages enable family ownership to improve decision-making efficiency, mitigate agency problems, and generate positive outcomes for minority shareholders (Anderson and Reeb, 2003; Maury, 2006). Despite this, the underlying motivations for managers to maintain cash reserves in family firms remain unclear. All raising important questions about family firm’s performance compared to non-family firms.

1.2 Problem Statement

Several studies have explored the influence of family ownership on firm performance, but no clear consensus has been reached (Villalonga and Amit, 2006; Hegde et al., 2020). A major challenge lies in the varying definitions of family firms, as highlighted by Harms (2014). Prior research suggests that family firms often leverage business opportunities, align their interests more effectively with shareholders, and secure loans on more favorable terms (Hegde et al., 2020). However, Anderson and Reeb (2003) argue that excessive family interference can shift the focus toward family interests rather than optimal firm performance, reducing decision-making efficiency and limiting external bidding opportunities. While empirical evidence has primarily focused on the United States of America (USA) (Anderson and Reeb, 2003; Fahlenbrach, 2009; Villalonga and Amit, 2006; Villalonga and Amit, 2009) and East Asia (Claessens et al., 2000; Z. Chen et al., 2005; Hegde et al., 2020), findings highlight regional differences in the impact of family ownership. In East Asia, weaker legal protections for minority shareholders, limited transparency, and concentrated family wealth hinder firm performance (Faccio and Lang, 2002; Claessens et al., 2000). Conversely, in Europe and the USA, stronger corporate governance frameworks mitigate these challenges, enabling family ownership to enhance decision-making efficiency, reduce agency problems, and benefit minority shareholders (Anderson and Reeb, 2003; Maury, 2006). These regional variations underscore the importance of regional specific characteristics of corporate governance in determining the effects of family ownership on firm performance. Studies reveal significant differences between the USA and Europe in corporate governance (Mendez, n.d.; Mintz, 2024; Ooghe and De Langhe, 2002). In the USA, dispersed ownership and a market-driven focus often limit family influence, emphasizing shareholder primacy and short-term goals (Villalonga and Amit, 2006; Anderson and Reeb, 2003). In contrast, Europe’s concentrated ownership structures and stakeholder-oriented governance enable family firms to exercise greater control, align long-term strategies, and enhance decision-making efficiency while balancing minority shareholder interests (Maury, 2006; Hegde et al., 2020). These differences highlight the role of regional governance frameworks and cultural factors in shaping the impact of family ownership on firm performance. Thus, understanding the relationship between family ownership and corporate governance requires a contextualized approach that accounts for regional variations. This study aims to bridge this gap by analyzing the contrasting effects of family ownership on firm performance across different governance systems and cultural settings.

1.3 Research Question

This study examines the differing influence of family ownership on firm performance in the USA and Europe, using the four most prominent corporate governance theories identified by Pindado and Requejo (2015). By employing linear regression analysis, it investigates performance variations between these family and non-family owned firms across the two regions. The following research question has been formulated to achieve this objective:

”Are there significant differences in the relationship between performance of family versus non-family firms between the USA and Europe?”

In the remainder of this thesis, the theoretical framework supporting the research question will be thoroughly explained, followed by the methodology used to conduct the study. Subsequently, the results will be presented. The paper will conclude with the findings, a discussion, and an exploration of the study’s limitations and recommendations for further research.

1.4 Contribution

This thesis contributes to the growing field of research on corporate governance by exploring the influence of family ownership on firm performance within a global context. Despite the significant presence of family firms worldwide, a clear gap exists in the literature regarding the relationship between regional corporate governance structures and the performance differences between family and non-family firms. By addressing these regional differences, this study seeks to bridge this gap and provide insights into how governance practices influence the effectiveness of family ownership in driving firm performance.

A key contribution lies in the comparison between the USA and Europe, two regions with distinct governance models. The USA’s market-driven framework, characterized by dispersed ownership and shareholder primacy, contrasts strongly with Europe’s concentrated ownership structures and stakeholder-oriented governance (Mendez, n.d.). This comparison sheds light on the role of corporate governance frameworks in enabling or constraining family influence and highlights how these differences impact firm performance.

Building on recent studies, this thesis incorporates updated empirical evidence to address the ongoing debate regarding the advantages and challenges of family ownership. It evaluates the regional variations in the effects of family governance, emphasizing how corporate governance frameworks shape outcomes for family and non-family firms.

Finally, this thesis draws on the four main theories discussed in the family-business literature, as outlined by Pindado and Requejo (2015). By integrating these theoretical perspectives, the research provides a comprehensive framework to assess the governance practices of family firms and their impact on performance across regions. This multi-theoretical approach allows for a nuanced understanding of the complexities surrounding family ownership and its interaction with corporate governance.

2 Theoretical Framework

2.1 Definition Family Firms

The definition of a family firm is a controversial issue, beginning with the debate over whether the founder must be the CEO or if it is sufficient for a relative of the founder to hold the position. McConaughy et al. (2001) & La Porta et al. (1999) propose a definition based on a minimum ownership threshold, such as 20%, arguing that higher ownership enhances the founder's influence over decision-making and corporate governance. Similarly, Anderson and Reeb (2003) highlight the nonlinear relationship between family ownership and firm performance, where performance improves with increased ownership but declines once ownership exceeds an optimal level of 32%. Berzins et al. (2018) define meaningful ownership as requiring a group of owners to hold more than half of the voting rights, though this definition is more applicable to private firms than public ones. Lower thresholds, such as the 10% proposed by Maury (2006), may be too inclusive, classifying too many firms as family-owned and weakening the observable "family effect". Arosa et al. (2010) incorporate the social dimension by defining family firms as those with a founder or family owner actively involved in monitoring the firm, though Berzins et al. (2018) critique this approach for its limitations. For instance, a family or founder with ownership but no formal governance role would not qualify as a family firm, while a family with a board seat but no ownership would. Therefore, the definition of Berzins et al. (2018), emphasizing the right to participate in governance rather than actual participation, is more robust. Ultimately, a high level of ownership that ensures the right to govern is the most critical factor in defining a family firm.

Following the research of Beene (2021) and Ulland and Upsaker (2020), we will, by combining Anderson and Reeb (2003), La Porta et al. (1999), and Berzins et al. (2018), define a family firm as a firm where the founder, or individuals related to the founder either by blood or marriage, has at least 20% ownership. Active participation by the family is not necessary, but rather a right.

2.2 Corporate Governance

This thesis focuses on the four most present ones according to Pindado and Requejo (2015).

2.2.1 Agency Theory

The first theory addressed by the researcher is agency theory, which is prominently discussed in family-firm business research, as evidenced by Pindado and Requejo (2015). Agency theory examines the relationship between a principal, an agent, and the contract binding them (Le Breton-Miller and Miller, 2009). It primarily focuses on potential conflicts arising from asymmetric information and misaligned interests between the two parties (Jensen and Meckling, 1979; Le Breton-Miller and Miller, 2009).

To mitigate these conflicts, principals can employ efficient incentives and monitoring mechanisms, which entail costs referred to as monitoring costs (Jensen and Meckling, 1979). Additionally, bonding costs arise when principals take measures to ensure that agents perform specific actions. If these actions are undertaken, the principal compensates the agent accordingly. Positive monitoring, such as encouraging desired behaviors, is also emphasized by Jensen and Meckling (1979). Despite these measures, welfare losses can occur when agents prioritize their own interests over those of the principal. Jensen and Meckling (1979) define agency costs as the sum of (1) monitoring costs for the principal, (2) bonding costs for the agent, and (3) welfare losses resulting from misaligned interests.

Family ownership can positively influence the principal-agent conflict, as the family's wealth is closely tied to the firm's success, providing strong incentives to monitor managers effectively (Anderson and Reeb, 2003; Le Breton-Miller and Miller, 2009; Sacristán-Navarro et al., 2011). Family firms also prioritize long-term survival and reputation, fostering enduring relationships with stakeholders (Sacristán-Navarro et al., 2011). However, these long-term commitments may bind family businesses to specific obligations.

Family-owned businesses often maintain long-term relationships with customers, suppliers, and partners, yet this focus may limit their willingness to explore new stakeholders who could offer more favorable contract terms (Le Breton-Miller and Miller, 2009). Additionally, family ownership is typically associated with less dispersed ownership, with families often taking on monitoring roles (Ang et al., 2000). Research suggests that in firms with less dispersed ownership, the likelihood of free-riding diminishes (Sacristán-Navarro et al., 2011).

Ang et al. (2000) explored two methods of measuring agency costs. The first method, the expense ratio, revealed that firms without a family or single owner holding more than 50% equity experienced significantly higher agency cost expenses. The second method provided insight into how effectively management allocated assets relative to sales.

The study highlighted that firms with dispersed ownership tend to under perform in asset utilization compared to firms managed by significant owners or families, with agency costs inversely related to the this ratio .

However, some studies challenge the concept that family ownership positively influences the agency relationship. Beyond the principal-agent conflict, there exists an "owner-owner conflict" stemming from misaligned interests between majority and minority shareholders (Le Breton-Miller and Miller, 2009). Minority shareholders often lack the contractual mechanisms or incentives to align the agent's interests with their own (Sacristán-Navarro et al., 2011). Yet, Anderson and Reeb (2003) argue that this issue is less significant in economies with robust shareholder protections. Furthermore, Hegde et al. (2020) contradict Le Breton-Miller and Miller (2009), asserting that the alignment of interests between the family and the firm benefits long-term minority shareholders.

Family firms often prioritize retaining control, even when an acquisition might be the most optimal strategic decision (Sacristán-Navarro et al., 2011). This tendency, combined with families' risk aversion, can constrain corporate innovation. Investments in business innovation, such as R&D, are perceived as risky, limiting such activities in family firms (Le Breton-Miller and Miller, 2009). A lack of innovation negatively impacts firm performance (Le Breton-Miller and Miller, 2009). Research supports this, with H.-L. Chen and Hsu (2009) acknowledging that family firms are less likely to make high-risk R&D investments. Patel and Chrisman (2014) further explain that family firms favor R&D projects that reliably enhance sales, avoiding speculative investments, even though such ventures could yield higher returns. Yet, R&D is essential for achieving competitive advantages and positive earnings (H.-L. Chen and Hsu, 2009). Block (2012) also highlights a negative correlation between family ownership and both R&D productivity and intensity.

Patel and Chrisman (2014) delve into why family firms approach R&D differently, emphasizing their focus on safeguarding socioemotional wealth and financial stability. Urbinati et al. (2017) reveal significant disparities in R&D investment levels between family and non-family firms. While family firms' low-risk strategies can be sustainable and effective, they may also limit innovation (Patel and Chrisman, 2014).

Another potential drawback of family ownership is the tendency to favor family members for executive roles, which can lead to overcommitment and the neglect of potentially better-qualified external candidates (Anderson and Reeb, 2003; Sacristán-Navarro et al., 2011; Le Breton-Miller and Miller, 2009).

2.2.2 Stewardship Theory

This research investigates stewardship theory as a contrasting perspective to agency theory. Stewardship theory suggests that individuals are not exclusively driven by self-interest but are also motivated to serve others (Le Breton-Miller and Miller, 2009). It emphasizes that individuals aim to fulfill higher-order needs, such as self-actualization, beyond economic incentives (Le Breton-Miller and Miller, 2009).

In family firms, stewardship theory suggests a focus on the long-term interests of the company and its stakeholders, given that the firm's future is inherently tied to the family's prospects, including family capital and career opportunities (Le Breton-Miller and Miller, 2009). Generosity among family firm managers often leads to improved firm performance, while involvement-oriented environments help reduce relationship conflicts (Kellermanns and Eddleston, 2006). An involvement-oriented management style is characterized by trust and informal agreements (Corbetta and Salvato, 2004). Kellermanns and Eddleston (2006) expand on this with the concept of a "participative strategy process," where family members actively engage in the operations and strategic planning of the firm. This approach fosters a better understanding of organizational goals and reduces individual biases (Kellermanns and Eddleston, 2006). Furthermore, the participative strategy process encourages individuals to share ideas and provide input, driving innovation in strategies that are critical to the firm's survival (Kellermanns and Eddleston, 2006). Such involvement also minimizes the likelihood of conflicts, as individuals feel valued and responsible for decisions (Kellermanns and Eddleston, 2006). While certain types of conflict can enhance performance, relationship conflicts are typically associated with stress and tension, which negatively affect firm performance (Kellermanns and Eddleston, 2006).

Altruistic behavior in family firms further contributes to reducing conflicts (Kellermanns and Eddleston, 2006). High levels of altruism strengthen family bonds, diminishing the reliance on authority-based decisions (Corbetta and Salvato, 2004). These bonds also promote shared values, encouraging a collective focus on organizational rather than personal goals (Corbetta and Salvato, 2004).

Additionally, the pro-organizational culture of family firms, centered on collectively maximizing firm performance, generates positive outcomes for both the family and the organization (Corbetta and Salvato, 2004).

2.2.3 Resource Based View

The resource-based view (RBV) suggests that companies derive competitive advantages from their unique combinations of assets and capabilities, collectively referred to as resources (Habbershon and Williams, 1999). These resources may be tangible or intangible, individual or organizational, and may also have a social dimension (Habbershon and Williams, 1999). When such resources are specific to a particular firm, they are termed idiosyncratic, offering potential competitive advantages (Habbershon and Williams, 1999). However, not all resources contribute to a competitive advantage. Tokarczyk et al. (2007) emphasize that only resources that are valuable and rare can lead to superior firm performance. Similarly, Cabrera-Suárez et al. (2011) argue that competitive advantage arises from a firm's ability to effectively utilize its resources and capabilities relative to competitors. To sustain this advantage, firms must ensure their resources remain isolated and protected from imitation or acquisition by competitors (Habbershon and Williams, 1999; Tokarczyk et al., 2007).

In family firms, the high commitment and dedication of family members serve as a strategic resource that can provide a competitive advantage (Cabrera-Suárez et al., 2011). This dedication often extends to relationships with customers, further enhancing the firm's competitive position (Cabrera-Suárez et al., 2011). Additionally, family firms may access more favorable banking terms due to their inside debt, which refers to financial obligations owed to insiders, which can also be viewed as an advantage (Hegde et al., 2020).

The unique combination of idiosyncratic resources and inimitable capabilities found in family firms, referred to as "familiness" can be a significant source of competitive advantage if leveraged effectively (Cabrera-Suárez et al., 2011; Habbershon and Williams, 1999; Sacristán-Navarro et al., 2011; Tokarczyk et al., 2007).

A crucial component of "familiness" is the tacit knowledge shared within the family (Cabrera-Suárez et al., 2011; Tokarczyk et al., 2007). Tacit knowledge is specialized, scarce, and difficult to exchange or imitate, providing the firm with unique opportunities (Cabrera-Suárez et al., 2011). However, as family members eventually leave the business, the transfer of this knowledge becomes critical (Cabrera-Suárez et al., 2011). In family firms, knowledge transfer is often more efficient than in non-family firms due to the close personal relationships between successors and predecessors (Cabrera-Suárez et al., 2011). This enhanced knowledge transfer capability strengthens the competitive advantage associated with "familiness" (Cabrera-Suárez et al., 2011).

2.2.4 Institutional Theory

Institutional theory focuses on the beliefs, values, practices, assumptions, and rules that establish a firm's credibility (Su et al., 2017). These norms and expectations aim to prevent undesirable behavior (Zhao et al., 2014). However, institutional expectations may not always align with economic rationality (Soleimanof et al., 2018). Family-firm owners, in particular, often face challenges arising from the contrasts between family norms and industry norms (Soleimanof et al., 2018). This conflict largely stems from the focus on socioemotional wealth (SEW), a concept emphasizing the emotional and non-financial value family members derive from their involvement in the business (Miller et al., 2013). Research by Miller et al. (2013) highlights that the level of family involvement significantly influences the importance placed on SEW.

Patel and Chrisman (2014) found that, while family and non-family firms may maintain similar levels of RD investment, family firms are better positioned to capitalize on these investments due to their SEW focus. In contrast, non-family firms often pursue exploratory investments, which increase variability in firm performance. Family firms, being more rooted in their historical foundations, are less likely to engage in experimental innovations (Patel and Chrisman, 2014). Instead, they prioritize industry-specific innovations, which enhance their SEW (Patel and Chrisman, 2014).

Family firms also interact differently with institutional contexts, the formal and informal frameworks within which organizations operate, compared to non-family firms. They share a common history that fosters unique values such as long-term orientation, legacy transmission, and loyalty (Parada et al., 2010; Soleimanof et al., 2018). Informal family meetings further encourage collaborative conversations, setting a shared vision for the firm (Soleimanof et al., 2018). However, this dynamic can also result in nepotism, the practice of favoring relatives or family members, where family institutions persistently dominate decision-making processes (Soleimanof et al., 2018).

Conformity, or adhering to industry standards, is another theme in family business literature. Miller et al. (2013) show that conformity correlates with the level of family involvement but is less common when the family-business founder is still active. Founders often embody entrepreneurial qualities, striving for innovation and pro-activeness rather than conformity (Miller et al., 2013).

Despite its challenges, conformity offers several advantages according to Miller et al.(2013). First, conforming strategies are market-tested, reducing business risks . Second, known demand allows for more accurate revenue and cost forecasting. Third, a conforming strategy can align with the achievement of SEW by enhancing community respect for the firm and the family. These benefits highlight how conformity, when balanced with innovation, can strengthen the performance and reputation of family firms.

2.3 Regional Difference

Corporate governance in the USA and the EU reflects the distinct economic, cultural, and regulatory landscapes of these regions, shaping how companies are governed and influencing their performance (Mendez, n.d.). This analysis integrates the four key theories as stated by Pindado and Requejo (2015), to highlight the contrasting corporate governance models in the USA and the EU and their relevance to family firms and their performance.

Agency theory focuses on reducing conflicts between owners and managers. In the USA, dispersed ownership is a characteristic of corporate governance, with institutional investors and publicly traded shares dominating ownership patterns (Pinto, 2010; Lessambo, 2014). This structure reduces direct owner control over management, relying instead on independent boards and regulatory frameworks to oversee and align managerial behavior (Lessambo, 2014). However, the separation of ownership and control often weakens monitoring and enables managerial opportunism, such as excessive compensation packages or short-term decision-making (Mendez, n.d.). In this context, family firms may hold a governance advantage by concentrating ownership and control, thereby reducing agency costs and enhancing oversight effectiveness.

In contrast, the EU is characterized by concentrated ownership, where significant shareholders, including family owners, exercise significant control. This reduces agency costs as families actively monitor management to align decisions with their long-term interests (Santos et al., 2014). However, the concentration of power can create "owner-owner conflicts," where majority shareholders prioritize their interests over those of minority shareholders (Le Breton-Miller and Miller, 2009). Additionally, the presence of employees on supervisory boards, as seen in Germany's two-tier governance system, can help balance these interests by incorporating diverse perspectives (Claassen, 2016). Family firms in the EU benefit from enhanced control and monitoring, however their advantage may be offset by concerns over minority shareholder protection and concentrated power.

Stewardship theory emphasizes trust, collaboration, and long-term orientation (Le Breton-Miller and Miller, 2009). In the USA, the unitary board system often combines the roles of CEO and board chairman, centralizing leadership to streamline decision-making (Pinto, 2010; Lessambo, 2014; Goh et al., 2014). While this structure can enhance efficiency, it risks excessive power concentration, potentially undermining board independence (Goh et al., 2014). The USA's market-driven governance model also prioritizes shareholder primacy, which can conflict with a family's focus on legacy and intergenerational goals (Lessambo, 2014).

The EU's governance model, particularly in countries with two-tier systems, promotes stewardship through the separation of supervisory and management boards (Claassen, 2016). Family members often serve on supervisory boards, directly influencing strategic direction while delegating day-to-day operations to professional managers (van Aaken et al., 2020). This structure aligns with the long-term orientation of family holdings, fostering trust and reducing conflicts. Additionally, employee representation on boards reinforces stewardship by integrating stakeholder interests into governance decisions (Claassen, 2016).

The RBV highlights the role of unique, firm-specific resources in achieving competitive advantage (Habbershon and Williams, 1999). In the USA, the governance model emphasizes market efficiency and innovation (Sun et al., 2019). Firms are encouraged to leverage resources dynamically, often prioritizing agility and adaptability to compete in fast-changing environments (Bavoso, 2016). This competitive market pressure may compel family firms to adopt more disciplined cost structures, professionalize management, and seek operational efficiencies. Additionally, U.S. family firms tend to exhibit higher levels of financial transparency and adopt performance-based compensation models that link managerial pay with performance metrics (Anderson and Reeb, 2003; Miller et al., 2007).

In contrast, the EU model supports the development and protection of idiosyncratic resources (Villalonga et al., 2015). Concentrated ownership and stakeholder-oriented practices allow family firms to capitalize on their unique assets, such as strong family branding, long-term stakeholder relationships, and industry-specific expertise (Habbershon and Williams, 1999). These strengths may translate more into reputational or market-based outcomes. Family firms in the EU may also be more risk-averse, emphasizing stability over profitability, further diluting profitability and performance metrics (Block, 2012).

Institutional theory examines the role of societal norms and regulations in shaping governance. In the USA,

institutional norms prioritize transparency, shareholder primacy, and market efficiency. This creates a predictable environment for investors, but it can conflict with the socioemotional priorities of family holdings, such as preserving legacy and community ties (Pinto, 2010; Lessambo, 2014). The unitary board structure, combined with dispersed ownership, often limits family influence unless families hold significant shares or board positions (Villalonga et al., 2015). U.S. institutional pressures often reward earnings growth and margin expansion, motivating family firms to optimize for accounting performance (Miller et al., 2007).

In the EU, institutional norms emphasize stakeholder engagement and long-term sustainability. Governance structures, such as the co-determination model in Germany, integrate family values with societal expectations, enhancing legitimacy and performance (Veldman, 2019; Gómez-Mejía et al., 2025). Family firms benefit from this alignment, as the regulatory environment supports their focus on legacy and intergenerational wealth. Nevertheless, this broader stakeholder focus may divert resources from immediate profitability goals, reducing short-term accounting performance (Gómez-Mejía et al., 2025).

2.4 Hypotheses

The literature review examined the four dominant corporate governance theories commonly applied in family business research, outlining their respective advantages and limitations as discussed in prior studies. Drawing on this theoretical foundation and empirical evidence on the performance implications of family ownership, the following hypotheses are proposed:

H1: Family-ownership has a significant positive influence on firm performance

H2: Family-ownership has a significant positive influence on firm performance in Europe

H3: Family-ownership has a significant positive influence on firm performance in the USA

Building further on the comparative analysis of corporate governance structures in the USA and EU, a final hypothesis addresses potential regional variation. Therefore the following hypothesis is formulated:

H4: The effect of family ownership on firm performance differs between Europe and the United States.

3 Methodology

The methodology section outlines the research approach and methods used to compare the performance of family and non-family firms across Europe and the USA.

3.1 Data Collection

Our sample will include family and non-family firms from Europe and the USA. We specifically use the database Orbis to extract information regarding public companies. Orbis is a comprehensive and powerful resource for data on private companies, while also covering publicly listed firms. A detailed overview of all exact variables extracted, the filters applied during the data extraction process and the exact extraction method can be found in Appendix 8.1. The final dataset resulted in a sample of approximately 30,000 companies, providing a robust foundation for the subsequent analysis.

3.2 Family Ownership Determination

To identify family firms within the dataset, we apply a multi-step classification process inspired by prior literature and adapted to the available ownership and management data in dataset. The final family firm label is assigned based on a combination of three filters. First, firms are flagged if a shareholder of type "one or more named individuals or families" hold over 20% direct or total ownership, following standard thresholds in family firm research. Second, we assess whether shareholder and board member names show high similarity using fuzzy string matching, capturing

active participation by the family. Third, shareholder names are checked for family-related keywords like “Family,” “Trust,” or “Holding”.

A firm is ultimately labeled as a family firm if it meets the ownership threshold and either the name-matching or keyword condition. This combined method ensures robust identification of family firms. Details are provided in Appendix 8.1.3.

3.3 Variables

3.3.1 Dependent Variables

The dependent variable of interest in this study is firm performance. Consistent with prior research (e.g., Anderson and Reeb, 2003; Barontini and Caprio, 2005; Hegde et al., 2020; Maury, 2006; Villalonga and Amit, 2006), firm performance will be assessed using both accounting and market-based measures. Specifically, we will evaluate firm performance using two approaches. The exact financial calculations can be found in Appendix 8.4.

The first measure of firm performance is Tobin’s Q, a widely used indicator of firm value. Tobin’s Q is calculated as the ratio of a firm’s market value to its assets (Villalonga and Amit, 2006; Barontini and Caprio, 2005). The Tobin’s Q is calculated by the market value of total assets divided by the book value of total assets. The market value of total assets is approximated by the market value of equity minus the total debt. In firms with multiple types of shares, some shares might not be publicly traded. To calculate the market value of common equity, this research, based on the approach of **villalonga how 20060** (2006), will multiply the total number of all outstanding shares by the price of the publicly traded shares.

The second method evaluates firm profitability using the return on assets (ROA) as the metric. Following Hegde et al. (2020), ROA is calculated by dividing the annual earnings before interest, taxes, depreciation, and amortization (EBITDA) by the average total assets. Essentially, ROA represents the ratio of a firm’s operating profit to its total assets (Barontini and Caprio, 2005).

Table 1: Dependent Variables

Variable	Measure	Source
<i>Tobin’s Q</i>	Market value divided by total assets	Barontini and Caprio, 2005; Villalonga and Amit, 2006
<i>Return on Assets (ROA)</i>	EBITDA divided by average total assets	Anderson and Reeb, 2003; Barontini and Caprio, 2005; Villalonga and Amit, 2006; Hegde et al., 2020

3.3.2 Independent Variables

In this research, family-ownership will be used as a dummy variable. Family-ownership will equal a “0” when the firm is a non-family firm and the variable will equal a “1” when the firm is a family-firm. How this research divided family and non-family firms is discussed in paragraph 3.2.

Table 2: Independent Variable

Variable	Measure	Source
<i>Family Firm</i>	Rated a ‘1’ if the founding family holds a fraction of the firm or is present on the board of directors. Rated a ‘0’ otherwise.	Anderson and Reeb, 2003; Patel and Chrisman, 2014; Villalonga and Amit, 2006

3.3.3 Control Variables

Since firm performance is influenced by multiple factors beyond the type of ownership, it is essential to include control variables in the analysis. As Maury (2006) highlights, industry and firm-specific characteristics play a significant role in shaping firm outcomes. The exact financial calculations can be found in Appendix 8.4.

3.3.3.1 Industry Variable

The first control variable is industry classification, based on grouping of two-digit SIC codes into 11 major divisions following the OSHA SIC Manual (of Labor, n.d.). While previous family firm studies typically control for industry using standard two-digit SIC codes (e.g., Anderson and Reeb, 2003; Barontini and Caprio, 2005; Hegde et al., 2020; Maury, 2006; Villalonga and Amit, 2006), this study uses a higher-level aggregation to improve interpretability and control for structural differences such as capital intensity, regulation, and market conditions.

3.3.3.2 Firm Age

The second control variable in this research is firm age, which may influence firm performance due to life cycle effects (Hegde et al., 2020). Firm age is measured by the difference between the founding year and the year 2024. The year 2024 is chosen since all financial data are measured at the end of year 2024. Older firms often benefit from accumulated goodwill, which can positively impact their performance (Hegde et al., 2020). This variable helps account for differences in firm maturity that could otherwise skew the analysis.

3.3.3.3 Firm Size

The third control variable is firm size, measured by calculating the natural logarithm of the book value of total assets. Firm size is frequently used in family-business research as it reflects the firm's life cycle stage and its potential influence on performance (Anderson and Reeb, 2003; Maury, 2006; vill). Larger firms may benefit from economies of scale, market power, and accumulated resources, which can positively impact performance (De Loecker et al., 2020; Eckert et al., 2022). Including this variable ensures size-related effects are controlled in the analysis.

3.3.3.4 Growth

The fourth control variable is firm growth, measured as the percentage change in net sales over the past three years, following the approach of Maury (2006). This metric reflects a firm's recent revenue expansion and serves as an indicator of strategic performance, competitiveness, and market dynamics.

3.3.3.5 Leverage

The last control variable is leverage as it reflects the company's financial risk and its potential to increase profits. To account for this, leverage will be measured by dividing the total book value of debt by the total book value of equity, following the methodologies used in previous research (Anderson and Reeb, 2003; Barontini and Caprio, 2005; Maury, 2006; Villalonga and Amit, 2006). Controlling for leverage ensures that the analysis considers the impact of financial risk on firm performance.

Table 3: Control Variables

Variable	Measure	Source
Industry	Industry is categorized by the method shown in Appendix 8.3, using the two-digit SIC code.	Beene, 2021; of Labor, n.d.; Anderson and Reeb, 2003; Barontini and Caprio, 2005; Hegde et al., 2020; Maury, 2006; Villalonga and Amit, 2006
Company Age	Difference between the founding year and the year 2024.	Beene, 2021; Anderson and Reeb, 2003; Barontini and Caprio, 2005; Hegde et al., 2020; Villalonga and Amit, 2006
Company Size	Natural log of the book value of total assets.	Beene, 2021; Anderson and Reeb, 2003; Barontini and Caprio, 2005; Hegde et al., 2020; Maury, 2006; Villalonga and Amit, 2006
Sales Growth	Average growth in net sales over the past three years.	Beene, 2021; Maury, 2006
Leverage	Total book value of debt divided by total book value of equity.	Beene, 2021; Anderson and Reeb, 2003; Barontini and Caprio, 2005; Villalonga and Amit, 2006

3.4 Sample

Financial corporations are excluded from the sample due to their distinct regulatory frameworks, which make them incomparable to non-financial corporations (Anderson and Reeb, 2003; Patel and Chrisman, 2014; Barontini and Caprio, 2005). Additionally, all companies with missing data are excluded to ensure the integrity of the analysis. Non-European Union countries, as well as the United Kingdom and Switzerland, are excluded due to differences in market structures and legal frameworks that could introduce bias. A complete list of included countries is provided in Appendix 8.2. Furthermore, the variables Tobin's Q, ROA, sales growth and leverage are winsorized at the 1st and 99th percentiles to mitigate the influence of extreme outliers. This results in a final data sample of 4026 firms, and an overview of the sample characteristics is shown in Appendix 8.6.

3.5 Regression Model

Our primary focus is on the potential difference of the potential relationship between family firms and firm performance in the USA and Europe. To thoroughly examine this relationship, we will incorporate multiple control variables into our analysis to ensure a robust assessment.

Since the dependent variable is measured in two forms, multiple regressions are conducted to assess the effects across both performance measures. The analysis is conducted using a regression model that includes both control variables and the independent variable. A full linear regression will be applied to the combined dataset to assess the overall effect of family ownership across both regions without separating them. Additionally this model will be applied to both Europe and the USA, allowing for a comparison of results across the two regions. By analyzing the differences, we aim to uncover regional variations in the relationship between family ownership and firm performance. Each of these models will also be tested for both dependent variables to check for significance in each context.

The regression approach suitable for our research is the Ordinary Least Squares (OLS) regression model, as employed by Barontini and Caprio (2005) and Beene (2021). Our dependent variables, Tobin's Q and ROA, capture firm performance, while our independent variable, family ownership, is represented by a dummy variable assigned a value of 1 for family firms. OLS is appropriate given the continuous nature of the dependent variables. OLS is well-suited for estimating linear relationships between continuous outcomes and categorical or continuous predictors, allowing for straightforward interpretation of the marginal effects of family ownership while controlling for other firm-level factors (Villalonga and Amit, 2006).

The general form of the regression model is:

$$\text{Firm Performance}_i = \alpha + \beta_1 \text{FamilyFirm}_i + \beta_2 \text{Industry}_i + \beta_3 \text{Age}_i + \beta_4 \text{Size}_i + \beta_5 \text{Growth}_i + \beta_6 \text{Leverage}_i + \varepsilon_i \quad (1)$$

4 Results

4.1 Descriptive Analysis

4.1.1 Descriptive Statistics Sample

Table 4 presents the descriptive statistics for the main variables used in the analysis.

The mean Tobin's Q is 1.433, indicating that, on average, firms in the sample are valued above their replacement cost of assets. However, the median is notably lower at 0.647, reflecting a skewed distribution driven by a few firms with very high values with a maximum of 18.178. Similarly, ROA has a mean close to zero with -0.0058 but a positive median of 0.088, suggesting most firms are modestly profitable, though some extreme negative outliers, with a minimum of -2.257 , pull down the average.

Firm age ranges widely, with an average of 32 years and a maximum of 266 years, indicating a diverse mix of young and very established firms. Firm size, measured as the natural logarithm of total assets, has a mean of 20.204, reflecting a distribution slightly right-skewed given the median of 20.492.

Sales growth over three years shows some variation, with a mean of 0.909 but a minimum of -1.000 and a maximum of 32.763, indicating the presence of both declining and expanding firms in the sample. Finally, leverage has a mean of 0.559 and a median of 0.299, highlighting that while many firms maintain moderate debt-to-equity ratios, some reach highly leveraged positions, with a maximum of 11.627, or even report negative values, possibly due to negative equity.

Table 4: Descriptive Statistics

Variable	Mean	Median	Min	Max	S.E.
<i>Tobin's Q</i>	1.43	0.65	-0.69	18.18	0.04
<i>ROA (%)</i>	-0.01	0.09	-2.26	0.39	0.01
<i>Firm Age</i>	32.28	24.00	0.00	266.00	0.48
<i>Firm Size</i>	20.20	20.49	2.30	27.21	0.04
<i>Sales Growth</i>	0.91	0.18	-1.00	32.76	0.06
<i>Leverage</i>	0.56	0.30	-9.35	11.63	0.03

Notes: This table provides a summary of the descriptive statistics for the dataset used in this study. The dataset consists of 4026 publicly listed firms from Europe and the United States, observed for the period 2021–2024. The European countries in the sample include: Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, and Sweden. Tobin's Q is measured as the market value divided by the total book value of the firm's assets. Return on Assets (ROA) is calculated as EBITDA divided by total assets. Firm Age is the difference between the year of incorporation and 2024. Firm Size is the natural logarithm of total assets. Sales Growth is the percentage change in net sales over three years. Leverage is total debt divided by total equity. For each variable, we report the mean, median, minimum, maximum, and standard error. The variables Tobin's Q, ROA, Sales Growth and Leverage are winsorized at the 1st and 99th percentiles.

4.1.2 Descriptive Statistics of Family- and Non-Family Firms

Table 5 presents the descriptive statistics comparing family firms ($N = 192$) and non-family firms ($N = 3,867$), along with t-tests for group differences.

On average, family firms show a slightly higher Tobin's Q of 1.512 compared to non-family firms with a Tobin's Q of 1.429, though the difference is not statistically significant. Similarly, average ROA is marginally higher for family firms with 1.4% versus non-family firms with -0.7% , but again this difference is not significant.

Firm age appears remarkably similar between the two groups, with family firms averaging 32.01 years and non-family firms averaging 32.30 years, suggesting no systematic age advantage. However, firm size reveals a significant difference. Family firms are notably smaller with a mean log assets of 19.04 compared to non-family firms with a mean log assets of 20.26, and this difference is highly significant.

Sales growth over three years is slightly lower for family firms with a mean of 0.848 compared to non-family firms with a mean of 0.912, but the difference is not statistically significant. Finally, leverage levels are comparable between groups, with family firms averaging 0.571 and non-family firms averaging 0.558.

Overall, the most notable and significant difference between family and non-family firms is their size, where family firms operate at a smaller scale. Other financial performance metrics, including market valuation, measured as Tobin's Q, profitability, measured as ROA, growth, and leverage, show no statistically significant differences between the groups.

Table 5: Descriptive Statistics of Family and Non-Family Firms

Variable	Family Firms (N = 192)					Non-Family Firms (N = 3867)					T-Statistic (p-value)
	Mean	Median	S.E.	Min	Max	Mean	Median	S.E.	Min	Max	
<i>Tobin's Q</i>	1.51	0.65	0.22	-0.69	18.18	1.43	0.65	0.04	-0.69	18.18	0.37 (0.713)
<i>ROA (%)</i>	0.01	0.09	0.02	-2.26	0.39	-0.01	0.09	0.01	-2.26	0.39	0.93 (0.355)
<i>Firm Age</i>	32.01	24.00	2.07	2.00	151.00	32.30	24.00	0.49	0.00	266.00	-0.14 (0.892)
<i>Firm Size</i>	19.04	19.17	0.19	11.36	26.38	20.26	20.56	0.05	2.30	27.21	-6.42 (8.84e-10)***
<i>Sales Growth</i>	0.85	0.17	0.31	-1.00	32.76	0.91	0.18	0.06	-1.00	32.76	-0.20 (0.842)
<i>Leverage</i>	0.57	0.25	0.14	-9.35	11.63	0.56	0.30	0.03	-9.35	11.63	0.09 (0.929)

Notes: *. Significant at the 0.1 level (2-tailed), **. Significant at the 0.05 level (2-tailed), ***. Significant at the 0.01 level (2-tailed). This table provides a summary of the descriptive statistics comparing family and non-family firms. The dataset consists of 4026 publicly listed firms from Europe and the United States, observed for the period 2021 to 2024. The European countries in the sample include: Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, and Sweden. Tobin's Q is measured as the market value divided by the total book value of the firm's assets. Return on Assets (ROA) is calculated as EBITDA divided by total assets. Firm Age is measured as the difference between the year of incorporation and the year 2024. Firm Size is the natural logarithm of total assets. Sales Growth is the percentage change in net sales over the past three years. Leverage is total debt divided by total equity. The variables Tobin's Q, ROA, Sales Growth and Leverage are winsorized at the 1st and 99th percentiles. For each variable, the mean, median, standard error, minimum, and maximum are reported, along with two-tailed t-tests comparing family and non-family firms.

4.1.3 Descriptive Statistics of Family- and Non-Family Firms in the EU

Table 6 reports the descriptive statistics comparing family firms (N = 107) and non-family firms (N = 1,411) within the EU sample, along with t-tests for group differences.

The average Tobin's Q is nearly identical across the two groups: 0.949 for family firms and 0.934 for non-family firms, with no significant difference. ROA also shows no meaningful distinction, averaging 8.5% for family firms and 8.0% for non-family firms.

Firm age is slightly lower for family firms, with a mean of 38.57 years, compared to non-family firms, with a mean of 43.50 years, but this difference is not statistically significant. The most notable and significant difference emerges in firm size. Family firms are considerably smaller, with a mean log assets of 18.66, than non-family firms, with a

mean log assets of 19.82, a difference that is highly significant.

Sales growth over three years shows no significant difference, with family firms averaging 0.49 and non-family firms slightly higher at 0.65. Similarly, leverage levels are comparable, averaging 0.50 for family firms and 0.55 for non-family firms.

Overall, the EU sample highlights that while family and non-family firms are largely similar across profitability, valuation, growth, and leverage metrics, they differ significantly in scale, with family firms operating on a notably smaller asset base.

Table 6: Descriptive Statistics of Family and Non-Family Firms in the EU

Variable	Family Firms (N = 107)					Non-Family Firms (N = 1411)					T-Statistic (p-value)
	Mean	Median	S.E.	Min	Max	Mean	Median	S.E.	Min	Max	
<i>Tobin's Q</i>	0.95	0.54	0.15	-0.69	10.87	0.93	0.46	0.05	-0.69	18.18	0.09 (0.927)
<i>ROA (%)</i>	0.09	0.12	0.02	-0.97	0.39	0.08	0.10	0.00	-2.26	0.39	0.32 (0.747)
<i>Firm Age</i>	38.57	26.50	3.10	2.00	151.00	43.50	31.00	1.00	2.00	266.00	-1.52 (0.132)
<i>Firm Size</i>	18.66	18.77	0.23	12.54	24.38	19.82	19.93	0.08	10.89	27.21	-4.80 (4.22e-06)***
<i>Sales Growth</i>	0.49	0.17	0.31	-0.97	32.76	0.65	0.15	0.08	-1.00	32.76	-0.48 (0.631)
<i>Leverage</i>	0.50	0.32	0.12	-6.81	3.73	0.55	0.34	0.04	-9.35	11.63	-0.37 (0.713)

Notes: *. Significant at the 0.1 level (2-tailed), **. Significant at the 0.05 level (2-tailed), ***. Significant at the 0.01 level (2-tailed). This table provides a summary of the descriptive statistics comparing family and non-family firms in the European Union. The dataset consists of **1503** publicly listed firms from Europe observed for the period 2021–2024. The European countries in the sample include: Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, and Sweden. Tobin's Q is measured as the market value divided by the total book value of the firm's assets. Return on Assets (ROA) is calculated as EBITDA divided by total assets. Firm Age is measured as the difference between the year of incorporation and 2024. Firm Size is the natural logarithm of total assets. Sales Growth is calculated as the percentage change in net sales over the past three years. Leverage is measured as total book value of debt divided by total book value of equity. The variables Tobin's Q, ROA, Sales Growth and Leverage are winsorized at the 1st and 99th percentiles. For each variable, the mean, median, standard error, minimum, and maximum are reported, as well as two-tailed t-tests comparing family and non-family firms.

4.1.4 Descriptive Statistics of Family- and Non-Family Firms in the USA

Table 7 reports the descriptive statistics comparing family firms (N = 85) and non-family firms (N = 2,456) within the USA sample, along with t-tests for group differences.

Family firms exhibit a higher average Tobin's Q of 2.223 compared to non-family firms, with a Tobin's Q of 1.712, though this difference is not statistically significant. ROA is slightly negative for family firms with a mean of -0.076 and also negative for non-family firms with a mean of -0.056, with no significant difference.

Firm age is lower among family firms, averaging 23.74 years compared to 25.88 years for non-family firms, but again, this difference does not reach significance. A key and significant distinction appears in firm size. Family firms are smaller with a mean log assets of 19.51 relative to non-family firms with a mean log assets of 20.51, with the

difference being statistically significant at the 1% level.

Sales growth over three years is positive for family firms with a mean of 1.30, and slightly lower for non-family firms with a mean of 1.06, although the difference is not significant. Leverage levels are comparable, with family firms averaging 0.661 and non-family firms 0.566.

Overall, the USA sample indicates that family firms differ most notably from non-family firms in their smaller size, while profitability, valuation, growth, and leverage metrics do not show statistically significant differences.

Table 7: Descriptive Statistics of Family and Non-Family Firms in the USA

Variable	Family Firms (N = 85)					Non-Family Firms (N = 2456)					T-Statistic (p-value)
	Mean	Median	S.E.	Min	Max	Mean	Median	S.E.	Min	Max	
<i>Tobin's Q</i>	2.22	0.87	0.45	-0.39	18.18	1.71	0.82	0.06	-0.69	18.18	1.12 (0.266)
<i>ROA (%)</i>	-0.08	0.06	0.04	-2.26	0.39	-0.06	0.08	0.01	-2.26	0.39	-0.46 (0.647)
<i>Firm Age</i>	23.74	18.00	2.28	2.00	105.00	25.88	20.00	0.47	0.00	124.00	-0.92 (0.359)
<i>Firm Size</i>	19.51	19.89	0.30	11.36	26.38	20.51	20.87	0.06	2.30	27.16	-3.32 (0.00131)***
<i>Sales Growth</i>	1.30	0.18	0.58	-1.00	32.76	1.06	0.20	0.09	-1.00	32.76	0.40 (0.688)
<i>Leverage</i>	0.66	0.09	0.27	-9.35	11.63	0.57	0.27	0.05	-9.35	11.63	0.35 (0.727)

Notes: *. Significant at the 0.1 level (2-tailed), **. Significant at the 0.05 level (2-tailed), ***. Significant at the 0.01 level (2-tailed). This table provides a summary of the descriptive statistics comparing family and non-family firms in the United States. The dataset consists of **2523** publicly listed firms observed for the period 2021–2024. Tobin's Q is measured as the market value divided by the total book value of the firm's assets. Return on Assets (ROA) is calculated as EBITDA divided by total assets. Firm Age is the difference between the year of incorporation and 2024. Firm Size is measured as the natural logarithm of total assets. Sales Growth is the percentage change in net sales over the past three years. Leverage is total book value of debt divided by total book value of equity. The variables Tobin's Q, ROA, Sales Growth and Leverage are winsorized at the 1st and 99th percentiles. For each variable, the mean, median, standard error, minimum, and maximum are reported, along with two-tailed t-tests comparing family and non-family firms.

The descriptive analysis of the full sample, as well as the breakdown by family and non-family firms and by region, reveals some variability across key variables. This variability underlines the importance of applying robust statistical techniques in the subsequent regression analyses. To mitigate the influence of extreme outliers, the variables Tobin's Q, ROA, Sales Growth and Leverage are winsorized at the 1st and 99th percentiles. This procedure reduces the risk of distortion caused by rare extreme observations. Additionally, further diagnostic checks were performed to assess skewness, heteroskedasticity, and multicollinearity, ensuring the validity and robustness of the regression results across all comparative analyses. The diagnostic tests indicate that the residuals in all models deviate significantly from normality as shown by the Shapiro-Wilk test with p-values below 0.001. In addition, the Breusch-Pagan test reveals strong evidence of heteroscedasticity in nearly all models with p-values below 0.01. Since both assumptions of classical linear regression are violated, robust standard errors of type HC3 have been applied. As a result, the confidence intervals and significance tests remain valid despite the violation of model assumptions.

4.2 Correlation Matrix

Table 8 presents the Pearson correlation coefficients between the key variables, indicating the strength and direction of their linear relationships.

The family firm variable shows only weak correlations with the independent and control variables, with the largest being a small negative correlation of -0.09 with firm size. This suggests that family ownership status is largely independent of firm profitability, market valuation, age, growth, and leverage in the sample, underscoring the need to test its effects directly through regression analysis rather than relying on simple bivariate relationships.

ROA shows a significant negative correlation with Tobin's Q with a correlation coefficient of -0.31 , suggesting that firms with higher market valuations relative to assets tend to have lower accounting profitability. ROA is positively and significantly correlated with firm size with a correlation coefficient of 0.50 , and firm age with a correlation coefficient of 0.20 , indicating that larger and older firms are generally more profitable.

Tobin's Q is negatively correlated with firm size with a correlation coefficient of -0.23 , and firm age with a correlation coefficient of -0.12 , suggesting that smaller and younger firms are often valued more highly by the market relative to their asset base.

Firm size and firm age are positively correlated with a correlation coefficient of 0.20 , reflecting the natural tendency of older firms to have accumulated more assets over time. Notably, growth over three years shows minimal correlation with other variables, except for a weak negative correlation with ROA with a correlation coefficient of -0.04 . Leverage shows small but significant correlations with Tobin's Q with a correlation coefficient of -0.03 and firm size with a correlation coefficient of 0.04 , though these relationships are weak in magnitude.

Overall, the matrix highlights expected relationships, such as the size-profitability link and the market preference for smaller, younger firms, while showing minimal ties between growth, leverage, or family ownership and the other firm characteristics. Importantly, the absence of strong multicollinearity between key independent variables confirms the suitability of applying linear regression models, as no excessive intercorrelations threaten the stability or interpretability of the regression estimates.

Table 8: Correlation Matrix

	Family Firm	ROA	Tobin's Q	Firm Size	Firm Age	Growth	Leverage
Family Firm	1.00***						
ROA	0.01	1.00***					
Tobin's Q	0.01	-0.31 ***	1.00***				
Firm Size	-0.09 ***	0.50 ***	-0.23 ***	1.00***			
Firm Age	-0.00	0.20 ***	-0.12 ***	0.20 ***	1.00***		
Growth	-0.00	-0.04 **	0.00	0.00	-0.01	1.00***	
Leverage	-0.00	0.02	-0.03 **	0.04 **	0.00	0.00	1.00***

Notes: *. Significant at the 0.1 level (2-tailed), **. Significant at the 0.05 level (2-tailed), ***. Significant at the 0.01 level (2-tailed). This table reports Pearson correlation coefficients between the key variables in the sample. Reported are two-tailed significance levels. *Family Firm* = family firm dummy, *ROA* = Return on Assets, *Tobin's Q* = market value over total assets, *Firm Size* = log total assets, *Growth* = three-year sales growth, *Leverage* = total debt over total equity. The variables Tobin's Q, ROA, Sales Growth and Leverage are winsorized at the 1st and 99th percentiles.

4.3 Linear Regression

4.3.1 Linear Regression

Table 9 presents the results of the OLS regressions estimating the effects of family firm status on Tobin's Q and ROA, including control variables and industry fixed effects.

For Tobin's Q, the family firm coefficient of -0.166 is negative but not statistically significant, suggesting that being a family firm does not significantly impact market valuation relative to assets. Among the control variables, firm size and firm age with coefficients of -0.177 and -0.006 are both negatively associated with Tobin's Q, indicating

that smaller and younger firms tend to have higher market valuations. Leverage also shows a small but significant negative effect of -0.074 , while growth has no significant effect.

For ROA, the family firm coefficient is positive and statistically significant with a regression coefficient of 0.099 , suggesting that family firms outperform non-family firms in terms of accounting profitability. Firm size with a coefficient of 0.062 and firm age with a coefficient of 0.001 are both positively associated with ROA, indicating that larger and older firms are more profitable. Growth shows a small but significant negative effect of -0.005 , while leverage does not show a significant effect on ROA.

The models achieve adjusted R^2 values of 0.080 for Tobin's Q and 0.277 for ROA, indicating modest explanatory power. Notably, the highest variance inflation factor (VIF) is low with a factor of 1.114 , suggesting no concerns with multicollinearity.

Overall, these regression results highlight that while family ownership or management does not significantly influence market-based performance, measured as Tobin's Q, it does positively affect accounting-based performance, measured as ROA, even after controlling for firm size, age, growth, leverage, and industry effects.

Table 9: Linear Regression

Variable	Tobin's Q	ROA
Intercept	4.7717 (0.6445)***	-1.2578 (0.0772)***
Family Firm	-0.1657 (0.2206)	0.0987 (0.0223)***
Firm Size	-0.1769 (0.0288)***	0.0621 (0.0035)***
Firm Age	-0.0055 (0.0009)***	0.0011 (0.0001)***
Sales Growth (3y)	0.0048 (0.0119)	-0.0046 (0.0020)**
Leverage	-0.0742 (0.0175)***	0.0010 (0.0017)
Industry Fixed Effects	Included	Included
Adjusted R^2	0.0795	0.2771
Highest VIF	1,114	1,114
Observations	4026	4026

Notes: *. Significant at the 0.1 level (2-tailed), **. Significant at the 0.05 level (2-tailed), ***. Significant at the 0.01 level (2-tailed). Robust standard errors in parentheses. This regression table shows the results of the OLS regressions estimating Tobin's Q and Return on Assets (ROA). The models include control variables. *Family Firm* is rated as '1' if the founding family holds a fraction of the firm or is present on the board of directors, and '0' otherwise. *Tobin's Q* is measured as the market value divided by the total assets of the firm. *ROA* is calculated by dividing EBITDA by total assets. *Company Age* is measured as the difference between the founding year and the year 2024. *Company Size* is measured as the natural logarithm of the book value of total assets. *Growth* is the average growth in net sales over the past three years. *Leverage* is total book debt over equity. *Industry* fixed effects are included based on previously discussed classifications. The variables Tobin's Q, ROA, Sales Growth and Leverage are winsorized at the 1st and 99th percentiles.

4.3.2 Linear Regression per Region

Table 10 presents the OLS regression results for Tobin's Q and ROA, estimated separately for the EU and USA samples, including control variables and industry fixed effects.

For the EU sample, the family firm coefficient of -0.093 is negative for Tobin's Q but not significant, indicating no measurable effect of family ownership on market valuation. For ROA, however, the family firm coefficient is positive and significant with a coefficient of 0.028 , suggesting a profitability advantage for family firms. Among controls, firm size negatively affects Tobin's Q with a coefficient of -0.051 and positively affects ROA with a coefficient of 0.018 , while firm age negatively affects Tobin's Q with a coefficient -0.004 and positively affects ROA with a coefficient 0.0002 . Leverage shows a significant negative effect on Tobin's Q with a coefficient -0.179 , but no effect on ROA.

For the USA sample, the family firm coefficient is positive but not significant for Tobin's Q with a coefficient 0.137 , indicating no clear market valuation advantage, but is positive and significant for ROA with a coefficient 0.085 , again suggesting higher profitability among family firms. Firm size has a significant negative effect on Tobin's

Q with a coefficient -0.305 and a positive effect on ROA with a coefficient 0.099 , while firm age positively affects ROA with a coefficient 0.0008 but not Tobin's Q. Leverage shows a significant negative effect on Tobin's Q with a coefficient -0.048 and no significant effect on ROA.

The adjusted R^2 values vary across models, with relatively low explanatory power for Tobin's Q, 0.080 for EU and 0.106 for the USA, but higher for ROA, especially in the USA, with 0.120 for EU and 0.433 for the USA. Variance inflation factors (VIFs) remain low across all models, indicating no multicollinearity issues.

Overall, these results suggest that family firms tend to achieve enhanced accounting-based performance, measured by ROA in both regions, but they do not enjoy a clear advantage in market-based valuation, measured by Tobin's Q. Firm size and age consistently play important roles in shaping both valuation and profitability outcomes across regions.

Table 10: Linear Regression Results per Region

Variable	Tobin's Q (EU)	ROA (EU)	Tobin's Q (USA)	ROA (USA)
Intercept	1.719 (0.345)***	-0.311 (0.047)***	7.865 (1.046)***	-2.056 (0.107)***
Family Firm	-0.093 (0.158)	0.028 (0.016)*	0.137 (0.438)	0.085 (0.039)**
Firm Size	-0.051 (0.016)***	0.018 (0.002)***	-0.305 (0.045)***	0.099 (0.005)***
Firm Age	-0.004 (0.001)***	0.0002 (0.0001)**	0.002 (0.002)	0.0008 (0.0002)***
Sales Growth (3y)	0.015 (0.016)	-0.004 (0.002)**	-0.005 (0.015)	-0.003 (0.002)
Leverage	-0.179 (0.062)***	0.004 (0.004)	-0.048 (0.018)***	-0.001 (0.002)
Industry Fixed Effects	Included	Included	Included	Included
Adjusted R^2	0.080	0.120	0.106	0.433
Highest VIF	1,175	1,175	1,141	1,141
Observations	1503	1503	2523	2523

Notes: ⁺. Significant at the 0.1 level (2-tailed), *. Significant at the 0.05 level (2-tailed), **. Significant at the 0.01 level (2-tailed), ***. Significant at the 0.001 level (2-tailed). Robust standard errors in parentheses. This table presents the results of ordinary least squares (OLS) regressions estimating Tobin's Q and Return on Assets (ROA) for family and non-family firms across two regional subsamples (EU and USA). *Family Firm* is coded '1' if the founding family holds a fraction of the firm or is represented on the board of directors, and '0' otherwise. *Tobin's Q* is the ratio of market value to total assets. *ROA* is calculated as EBITDA over total assets. *Firm Size* is measured as the natural logarithm of total assets, and *Firm Age* as the difference between incorporation year and 2024. *Growth* captures average net sales growth over the past three years, and *Leverage* is total debt divided by equity. Industry fixed effects are included. The variables Tobin's Q, ROA, Sales Growth and Leverage are winsorized at the 1st and 99th percentiles.

The results show that family ownership is significantly positively associated with ROA in both regions, but with a notably stronger effect in the United States. In the U.S. sample, the regression coefficient for family ownership on ROA is 0.085 , compared to a smaller but still significant coefficient of 0.028 in the European sample. This indicates that, although family ownership enhances profitability in both contexts, the effect is more pronounced in the U.S. firms.

In contrast, no significant effect of family ownership is observed on Tobin's Q in either region. This suggests that market valuation does not systematically differ between family and non-family firms in Europe or the United States, despite the observed differences in accounting performance.

4.4 Potential Endogeneity

This research must be mindful of the potential risk of endogeneity. As Anderson and Reeb (2003) highlight, families may retain ownership precisely because the firm is performing well. Additionally, families may benefit from their privileged access to company-specific information due to their involvement in the firm (Anderson and Reeb, 2003; Barontini and Caprio, 2005; Maury, 2006; Villalonga and Amit, 2006). This informational advantage may allow families to remain involved only when the firm's prospects are favorable (Anderson and Reeb, 2003; Barontini and

Caprio, 2005; Maury, 2006; Villalonga and Amit, 2006), and conversely, they may exit when future expectations are poor (Villalonga and Amit, 2006).

Despite these considerations, prior literature suggests that the results of our regression analysis do not necessarily have to be interpreted as endogenous (Anderson and Reeb, 2003; Barontini and Caprio, 2005). However, as Anderson and Reeb (2003) note, the possibility of endogeneity cannot be entirely ruled out.

5 Conclusion

5.1 Hypothesis Testing

This thesis set out to investigate whether there are significant differences in the relationship between firm performance and family ownership across regions, specifically between the United States and Europe. Based on regression analyses of 4,026 publicly listed firms over the period 2021–2024, the findings offer partial support for the stated hypotheses based on the regression analyses presented in Table 10.

Hypothesis 1: *Family-ownership has a significant positive influence on firm performance*

The evidence for hypotheses 1 is mixed. While family ownership is positively and significantly associated with accounting-based performance (ROA), no significant effect is found for market-based performance (Tobin's Q). This suggests that family ownership is linked to higher profitability, but this advantage is not reflected in market valuation. Therefore, hypotheses 1 is only partially supported.

Hypothesis 2: *Family-ownership has a significant positive influence on firm performance in Europe*

For hypotheses 2 the evidence is mixed. While the regression results show a positive and statistically significant effect of family ownership on accounting-based performance, measured as ROA, no significant effect is observed for market-based performance, measures as Tobin's Q. Therefore, hypotheses 2 is partially supported. Family ownership positively influences profitability but does not significantly impact market valuation in the European sample.

Hypothesis 3: *Family-ownership has a significant positive influence on firm performance in the United States*

For hypotheses 3 the results similarly show a significant positive effect on ROA but no significant effect on Tobin's Q. Thus, hypotheses 3 is also partially supported. Family ownership improves accounting-based performance but does not translate into a higher market valuation among U.S. firms.

Hypothesis 4: *The effect of family ownership on firm performance differs between Europe and the United States*

Regarding Hypothesis 4, the results provide partial support. Specifically, while both regions exhibit a significant positive association between family ownership and accounting performance, measured as ROA, the effect appears stronger in the U.S., with a regression coefficient of 0.085 compared to 0.028 in Europe. However, no significant relationship is observed for market-based performance, measured as Tobin's Q in either region.

These findings suggest that the performance premium associated with family ownership is not universally stronger in Europe, as often theorized. On the contrary, the U.S. sample shows a slightly more pronounced positive effect on profitability. Thus, Hypothesis 4 is supported in terms of the existence of regional differences.

In summary, family ownership is positively related to firm profitability, measured as ROA, in both regions. However, this advantage is more pronounced in the United States. Despite this, it does not translate into higher market valuation, as measured by Tobin's Q, in either region.

5.2 Interpretation

The larger effect of family ownership on ROA in the United States challenges the assumption that Europe's stakeholder-oriented governance and institutional frameworks better support the advantages of family control. Several factors could explain this outcome.

First, the resource-based view suggests that superior firm performance stems from the effective deployment of unique, inimitable resources, often described in the context of "familiness" (Habbershon and Williams, 1999; Cabrera-Suárez et al., 2011). In the United States, family firms operate in a highly competitive and efficiency-driven market environment, which may push them to leverage their internal resources, such as tacit knowledge, reputation, and relational contracts, more aggressively to survive and outperform peers (Carney, 2005; Miller et al., 2007). As a result, they may be more likely to translate "familiness" into tangible financial outcomes.

Second, from an agency theory perspective, U.S. family firms may experience lower agency costs due to more clearly aligned incentives between owners and managers. In environments with dispersed ownership, characteristic of the U.S., the presence of a controlling family can act as a valuable governance mechanism, reducing managerial

opportunism and enhancing profitability (Anderson and Reeb, 2003). In contrast, European firms often already operate under concentrated ownership structures, meaning that the incremental governance benefit provided by family ownership may be relatively smaller (Santos et al., 2014).

Third, U.S. family firms may be less risk-averse in financial decision-making than their European counterparts. European family firms are typically embedded in institutional contexts that prioritize continuity, conformity, and the preservation of socioemotional wealth, which often results in more conservative investment behavior, particularly in areas related to innovation and growth (Gomez-Mejia et al., 2007; Patel and Chrisman, 2014). While this can promote stability, it may also limit financial performance, particularly in fast-moving industries.

The absence of a significant relationship between family ownership and Tobin's Q in either region suggests that financial markets may not fully price in the internal advantages of family firms. Tobin's Q is sensitive to investor perceptions, and the lack of a clear premium may reflect skepticism regarding succession risks, or underinvestment in growth opportunities, issues commonly associated with family governance structures (Block, 2012; Villalonga et al., 2015). Additionally, the market may undervalue long-term relational advantages that are not immediately reflected in quarterly earnings or innovation output.

5.3 Implications

This research contributes to the family firm literature by providing cross-regional evidence on the performance effects of family ownership. While prior studies have debated whether family firms outperform non-family firms, particularly in European institutional contexts, this study shows that the positive influence on accounting-based performance (ROA) is present in both Europe and the USA. Importantly, the absence of a significant effect on market-based performance (Tobin's Q) suggests that investors may not fully price in the profitability advantage of family firms or that other market factors offset this advantage.

Furthermore, the finding that the positive performance effect is not stronger in Europe challenges commonly held assumptions about the importance of institutional or cultural contexts in amplifying the family firm advantage.

For practitioners, the findings highlight that family ownership can be a meaningful driver of internal firm performance in terms of profitability but does not necessarily translate into higher market valuation. This is valuable for family business owners and managers, as it suggests that capital markets may undervalue the operational strengths of family-controlled firms, potentially creating opportunities to improve investor communication or corporate governance practices to better align market perceptions with internal performance.

For investors, the results indicate that relying solely on market metrics like Tobin's Q may overlook important dimensions of family firm performance, such as sustained profitability. This opens up opportunities for more nuanced investment strategies that factor in family ownership structures when assessing firm value.

Finally, for policymakers, the study suggests that the advantages of family ownership are region-specific, with stronger effects observed in the USA. This implies that supportive family business policies or governance frameworks may need to be tailored to regional institutional contexts rather than assuming universal applicability.

In sum, family ownership is associated with improved internal performance across both regions, but the strength and visibility of this effect depend on regional institutional dynamics and market perceptions. These findings challenge prevailing assumptions and invite a more nuanced understanding of the family firm advantage in global capital markets.

6 Discussion

6.1 Limitations

While this research provides valuable insights into the relationship between family ownership and firm performance, several limitations should be acknowledged.

First, the study relies on cross-sectional data, which limits the ability to establish causal relationships. Although the regressions control for key firm characteristics and industry effects, potential endogeneity, such as reverse causality or omitted variable bias, cannot be fully ruled out, as noted in prior literature (Anderson and Reeb, 2003). Future research using panel data or instrumental variable approaches could provide stronger causal inference.

Second, the study measures family ownership using available database indicators and classifications, which may not fully capture the complexity and heterogeneity of family involvement. Factors such as generational stage, family governance structures, or the active versus passive role of family members were not incorporated, yet these could meaningfully influence performance outcomes.

Third, the analysis is limited to publicly listed firms, which may bias the results toward larger, more transparent companies. Many family businesses operate privately, and their governance dynamics or performance characteristics may differ substantially from those observed in the public domain.

Fourth, the analysis focuses on two specific regions, Europe and the USA, which, while offering valuable comparative insights, may limit the generalizability of the findings to other contexts such as Asia, Latin America, or emerging markets, where family businesses may operate under very different institutional or cultural conditions.

Fifth, the performance measures used, ROA and Tobin's Q, capture only certain dimensions of firm success. Non-financial outcomes, such as employee satisfaction, innovation, or long-term resilience, are not assessed but may be particularly relevant in the context of family firms.

Finally, although the models include controls for firm size, age, growth, leverage, and industry effects, other potentially important variables, such as corporate governance, board composition, or market conditions, were not included due to data limitations. This raises the possibility of omitted variable bias that could affect the interpretation of the results.

6.2 Future Research

Building on the findings and limitations of this study, several promising directions for future research emerge.

First, future studies could employ panel data to better capture the dynamic effects of family ownership on firm performance over time. Longitudinal designs would allow researchers to address causality concerns more rigorously and examine how the relationship between family ownership and performance evolves across firm life cycles, generational transitions, or external shocks.

Second, it would be valuable to incorporate more detailed measures of family involvement. Future research could distinguish between founder-led firms, second-generation, and later-generation family firms, or examine the role of active family management versus passive ownership. Such distinctions could shed light on whether the performance effects identified here are uniform across all family firms or concentrated in specific subgroups.

Third, expanding the geographic scope to include emerging markets or non-Western contexts could offer richer comparative insights. Given the strong cultural and institutional differences in how family firms operate globally, studies that include regions like Asia, Latin America, or Africa could test the generalizability of the patterns observed in this study and explore region-specific mechanisms.

Fourth, future work could move beyond financial performance measures to explore non-financial outcomes, such as innovation, sustainability practices, employee well-being, or reputation. Family firms are often characterized by long-term orientations and socioemotional goals, which may not always be reflected in short-term financial indicators.

Finally, there is room for research exploring investor perceptions and market mechanisms. Given the finding that market-based valuations do not fully reflect the profitability advantages of family firms, future studies could investigate how information asymmetry, disclosure practices, or investor biases influence the market's assessment of family firms.

7 References

References

- Anderson, R. C., & Reeb, D. M. (2003). Founding-Family Ownership and Firm Performance: Evidence from the S&P 500 [Publisher: [American Finance Association, Wiley]]. *The Journal of Finance*, 58(3), 1301–1328. Retrieved January 17, 2025, from <https://www.jstor.org/stable/3094581>
- Ang, J. S., Cole, R. A., & Lin, J. W. (2000). Agency Costs and Ownership Structure. *The Journal of Finance*, 55(1), 81–106. <https://doi.org/10.1111/0022-1082.00201>
- Arosa, B., Iturralde, T., & Maseda, A. (2010). Ownership structure and firm performance in non-listed firms: Evidence from Spain. *Journal of Family Business Strategy*, 1(2), 88–96. <https://doi.org/10.1016/j.jfbs.2010.03.001>
- Barontini, R., & Caprio, L. (2005, June). The Effect of Family Control on Firm Value and Performance. Evidence from Continental Europe. <https://doi.org/10.2139/ssrn.675983>
- Bavoso, V. (2016). Corporate Governance in Developing and Emerging Markets: Debates, Models and New Institutional Economics, edited by F. Ngwu, retrieved January 18, 2025, from <https://consensus.app/papers/corporate-governance-in-developing-and-emerging-markets-bavoso/d67735f31e495f1f9a0b732b59d7d41f/>
- Beene, S. (2021, July). The relationship between family-ownership and firm performance [Publisher: University of Twente]. Retrieved January 17, 2025, from <https://essay.utwente.nl/87704/>
- Berzins, J., Böhren, Ø., & Stacescu, B. (2018, June). The Governance and Finance of Norwegian Family Firms: Main Characteristics of the Population. <https://doi.org/10.2139/ssrn.3182411>
- Block, J. H. (2012). R&D investments in family and founder firms: An agency perspective. *Journal of Business Venturing*, 27(2), 248–265. <https://doi.org/10.1016/j.jbusvent.2010.09.003>
- Bunkanwanicha, P., Wiwattanakantang, Y., & Fan, P. H. J. (2013). The Value of Marriage to Family Firms. *Journal of Financial and Quantitative Analysis*. <https://doi.org/10.1017/S0022109013000148>
- Cabrera-Suárez, M. K., de la Cruz Déniz-Déniz, M., & Martín-Santana, J. D. (2011). Familiness and market orientation: A stakeholder approach. *Journal of Family Business Strategy*, 2(1), 34–42. <https://doi.org/10.1016/j.jfbs.2011.01.001>
- Carney, M. (2005). Corporate Governance and Competitive Advantage in Family-Controlled Firms. *Entrepreneurship Theory and Practice*, 29, 249–265. <https://doi.org/10.1111/j.1540-6520.2005.00081.x>
- Chen, E.-T., Gray, S., & Nowland, J. (2013). Family Representatives in Family Firms [eprint: <https://onlinelibrary.wiley.com/doi/10.1111/corg.12009>]. *Corporate Governance: An International Review*, 21(3), 242–263. <https://doi.org/10.1111/corg.12009>
- Chen, H.-L., & Hsu, W.-T. (2009). Family Ownership, Board Independence, and R&D Investment. *Family Business Review*, 22(4), 347–362. <https://doi.org/10.1177/0894486509341062>
- Chen, Z., Cheung, Y.-L., Stouraitis, A., & Wong, A. W. S. (2005). Ownership concentration, firm performance, and dividend policy in Hong Kong. *Pacific-Basin Finance Journal*, 13(4), 431–449. <https://doi.org/10.1016/j.pacfin.2004.12.001>
- Chrisman, J. J., Chua, J. H., Pearson, A. W., & Barnett, T. (2012). Family Involvement, Family Influence, and Family-Centered Non-Economic Goals in Small Firms - James J. Chrisman, Jess H. Chua, Allison W. Pearson, Tim Barnett, 2012. *Entrepreneurship Theory and Practice*, Pages: 267 - 293 (Volume 36, Issue 2). Retrieved December 18, 2024, from https://journals.sagepub.com/doi/full/10.1111/j.1540-6520.2010.00407.x?casa_token=fJ23zIWshHaMAAAAA%3AplOosjenT6cCqNzByRv_foYryHg_dSh-nHrTXQyn3dyeFFgz5GFN9v8aU1ouSarz7fbo
- Chrisman, J. J., Sharma, P., Steier, L. P., & Chua, J. H. (2013). The Influence of Family Goals, Governance, and Resources on Firm Outcomes [Publisher: SAGE Publications Inc]. *Entrepreneurship Theory and Practice*, 37(6), 1249–1261. <https://doi.org/10.1111/etap.12064>
- Claassen, D. (2016, June). Supervisory Board Compensation Structure, Employee Representation, and Earnings Management: Evidence from German Boards. <https://doi.org/10.2139/ssrn.2482487>
- Claessens, S., Djankov, S., & Lang, L. H. P. (2000). The separation of ownership and control in East Asian Corporations. *Journal of Financial Economics*, 58(1), 81–112. [https://doi.org/10.1016/S0304-405X\(00\)00067-2](https://doi.org/10.1016/S0304-405X(00)00067-2)
- Corbetta, G., & Salvato, C. A. (2004). The Board of Directors in Family Firms: One Size Fits All? [Publisher: SAGE Publications Inc]. *Family Business Review*, 17(2), 119–134. <https://doi.org/10.1111/j.1741-6248.2004.00008.x>
- De Loecker, J., Eeckhout, J., & Unger, G. (2020). The Rise of Market Power and the Macroeconomic Implications*. *The Quarterly Journal of Economics*, 135(2), 561–644. <https://doi.org/10.1093/qje/qjz041>

- Eckert, S., Koppe, M., Burkatzki, E., Eichtopf, S., & Scharf, C. (2022). Economies of Scale: The Rationale Behind the Multinationality-Performance Enigma. *Management International Review*, 62(5), 681–710. <https://doi.org/10.1007/s11575-022-00473-2>
- Elstrodt, H. P., & Poulet, J. M. (2014). Perspectives on founder- and family-owned businesses. *McKinsey, Quart 4-14*. Retrieved December 18, 2024, from <https://www.mckinsey.com/industries/private-capital/our-insights/perspectives-on-founder-and-family-owned-businesses>
- Faccio, M., & Lang, L. H. P. (2002). The ultimate ownership of Western European corporations. *Journal of Financial Economics*, 65(3), 365–395. [https://doi.org/10.1016/S0304-405X\(02\)00146-0](https://doi.org/10.1016/S0304-405X(02)00146-0)
- Fahlenbrach, R. (2009). Founder-CEOs, Investment Decisions, and Stock Market Performance [Publisher: [Cambridge University Press, University of Washington School of Business Administration]]. *The Journal of Financial and Quantitative Analysis*, 44(2), 439–466. Retrieved December 18, 2024, from <https://www.jstor.org/stable/40505931>
- Goh, C. F., Rasli, A., & Khan, S.-U.-R. (2014). CEO duality, board independence, corporate governance and firm performance in family firms: Evidence from the manufacturing industry in Malaysia. *Asian Business & Management*, 13(4), 333–357. <https://doi.org/10.1057/abm.2014.4>
- Gomez-Mejia, L., Haynes, K. T., Núñez-Nickel, M., Jacobson, K. J. L., & Moyano-Fuentes, J. (2007). Socioemotional wealth and business risks in family-controlled firms: Evidence from Spanish olive oil mills. *Administrative Science Quarterly*, 52(1), 106–137. <https://doi.org/10.2189/asqu.52.1.106>
- Gómez-Mejía, L. R., Muñoz-Bullón, F., Requejo, I., & Sanchez-Bueno, M. J. (2025). Ethical Correlates of Family Control: Socioemotional Wealth, Environmental Performance, and Financial Returns. *Journal of Business Ethics*, 198(4), 893–917. <https://doi.org/10.1007/s10551-025-05943-9>
- Habbershon, T. G., & Williams, M. L. (1999). A Resource-Based Framework for Assessing the Strategic Advantages of Family Firms [Publisher: SAGE Publications Inc]. *Family Business Review*, 12(1), 1–25. <https://doi.org/10.1111/j.1741-6248.1999.00001.x>
- Harms, H. (2014). Review of Family Business Definitions: Cluster Approach and Implications of Heterogeneous Application for Family Business Research [Number: 3 Publisher: Multidisciplinary Digital Publishing Institute]. *International Journal of Financial Studies*, 2(3), 280–314. <https://doi.org/10.3390/ijfs2030280>
- Hegde, S., Seth, R., & Vishwanatha, S. R. (2020). Ownership Concentration and Stock Returns: Evidence from Family Firms in India [Publisher: Elsevier]. *Pacific-Basin Finance Journal*, 61, 101330. <https://doi.org/10.1016/j.pacfin.2020.101330>
- Jensen, M. C., & Meckling, W. H. (1979). Rights and Production Functions: An Application to Labor-Managed Firms and Codetermination [Publisher: University of Chicago Press]. *The Journal of Business*, 52(4), 469–506. Retrieved January 17, 2025, from <https://www.jstor.org/stable/2352442>
- Kellermanns, F. W., & Eddleston, K. A. (2006). Corporate Entrepreneurship in Family Firms: A Family Perspective [Publisher: SAGE Publications Inc]. *Entrepreneurship Theory and Practice*, 30(6), 809–830. <https://doi.org/10.1111/j.1540-6520.2006.00153.x>
- Khanna, T., & Rivkin, J. W. (2001). Estimating the Performance Effects of Business Groups in Emerging Markets [Publisher: Wiley]. *Strategic Management Journal*, 22(1), 45–74. Retrieved December 18, 2024, from <https://www.jstor.org/stable/3094254>
- Kuan, T.-H., Li, C.-S., & Chu, S.-H. (2011). Cash holdings and corporate governance in family-controlled firms. *Journal of Business Research*, 64(7), 757–764. <https://doi.org/10.1016/j.jbusres.2010.07.004>
- La Porta, R., Lopez-de-Silanes, F., & Shleifer, A. (1999). Corporate Ownership around the World [Publisher: [American Finance Association, Wiley]]. *The Journal of Finance*, 54(2), 471–517. Retrieved January 17, 2025, from <https://www.jstor.org/stable/2697717>
- Le Breton-Miller, I., & Miller, D. (2009). Agency vs. Stewardship in Public Family Firms: A Social Embeddedness Reconciliation [Publisher: SAGE Publications Inc]. *Entrepreneurship Theory and Practice*, 33(6), 1169–1191. <https://doi.org/10.1111/j.1540-6520.2009.00339.x>
- Lessambo, F. I. (2014). Corporate Governance in the United States of America. In F. I. Lessambo (Ed.), *The International Corporate Governance System: Audit Roles and Board Oversight* (pp. 46–80). Palgrave Macmillan UK. https://doi.org/10.1057/9781137360014_5
- Liu, Q., Luo, T., & Tian, G. G. (2015). Family control and corporate cash holdings: Evidence from China. *Journal of Corporate Finance*, 31, 220–245. <https://doi.org/10.1016/j.jcorpfin.2015.02.007>

- Liu, Y. (2011). FOUNDING FAMILY OWNERSHIP AND CASH HOLDINGS [Publisher: Southern Finance Association]. *Journal of Financial Research*, 34(2), 279–294. Retrieved December 18, 2024, from https://econpapers.repec.org/article/blajfnres/v_3a34_3ay_3a2011_3ai_3a2_3ap_3a279-294.htm
- Madison, K., Holt, D. T., Kellermanns, F. W., & Ranft, A. L. (2016). Viewing Family Firm Behavior and Governance Through the Lens of Agency and Stewardship Theories [Publisher: SAGE Publications Inc]. *Family Business Review*, 29(1), 65–93. <https://doi.org/10.1177/0894486515594292>
- Maury, B. (2006). Family ownership and firm performance: Empirical evidence from Western European corporations. *Journal of Corporate Finance*, 12(2), 321–341. <https://doi.org/10.1016/j.jcorpfin.2005.02.002>
- Mcconaughey, D. L., Matthews, C. H., & Fialko, A. S. (2001). Founding Family Controlled Firms: Performance, Risk, and Value. *Journal of Small Business Management*, 39(1), 31–49. <https://doi.org/10.1111/0447-2778.00004>
- Mendez, M. A. (n.d.). Corporate Governance A US/EU Comparison. *Global Business Center, University of Washington*.
- Miller, D., Breton-Miller, I. L., & Lester, R. H. (2013). Family Firm Governance, Strategic Conformity, and Performance: Institutional vs. Strategic Perspectives. *Organization Science*, 24(1), 189–209. <https://doi.org/10.1287/orsc.1110.0728>
- Miller, D., Le Breton-Miller, I., Lester, R. H., & Cannella, A. A. (2007). Are family firms really superior performers? *Journal of Corporate Finance*, 13(5), 829–858. <https://doi.org/10.1016/j.jcorpfin.2007.03.004>
- Mintz, S. M. (2024). A comparison of corporate governance systems in the U.S., UK and Germany. *Corporate Ownership & Control / Volume 3, Issue 4, Summer 2006*. <https://doi.org/10.22495/cocv3i4p12>
- of Labor, U. S. D. (n.d.). SIC Manual. Retrieved June 4, 2025, from <https://www.osha.gov/data/sic-manual>
- Ooghe, H., & De Langhe, T. (2002). The Anglo-American versus the Continental European corporate governance model: Empirical evidence of board composition in Belgium [Publisher: MCB UP Ltd]. *European Business Review*, 14(6), 437–449. <https://doi.org/10.1108/09555340210448794>
- Ozkan, A., & Ozkan, N. (2004). Corporate cash holdings: An empirical investigation of UK companies. *Journal of Banking & Finance*, 28(9), 2103–2134. <https://doi.org/10.1016/j.jbankfin.2003.08.003>
- Parada, M. J., Nordqvist, M., & Gimeno, A. (2010). Institutionalizing the Family Business: The Role of Professional Associations in Fostering a Change of Values [Publisher: SAGE Publications Inc]. *Family Business Review*, 23(4), 355–372. <https://doi.org/10.1177/0894486510381756>
- Patel, P. C., & Chrisman, J. J. (2014). Risk Abatement as a Strategy for R&d Investments in Family Firms [Publisher: Wiley]. *Strategic Management Journal*, 35(4), 617–627. Retrieved January 17, 2025, from <https://www.jstor.org/stable/24037181>
- Pindado, J., & Requejo, I. (2015). Family Business Performance from a Governance Perspective: A Review of Empirical Research [eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/ijmr.12040>]. *International Journal of Management Reviews*, 17(3), 279–311. <https://doi.org/10.1111/ijmr.12040>
- Pinto, A. R. (2010). An Overview of United States Corporate Governance in Publicly Traded Corporations [Publisher: Oxford University Press]. *The American Journal of Comparative Law*, 58, 257–283. Retrieved January 18, 2025, from <https://www.jstor.org/stable/20744541>
- Sacristán-Navarro, M., Gómez-Ansón, S., & Cabeza-García, L. (2011). Large shareholders' combinations in family firms: Prevalence and performance effects. *Journal of Family Business Strategy*, 2(2), 101–112. <https://doi.org/10.1016/j.jfbs.2011.03.001>
- Santos, M. S., Moreira, A. C., & Vieira, E. S. (2014). Ownership concentration, contestability, family firms, and capital structure. *Journal of Management & Governance*, 18(4), 1063–1107. <https://doi.org/10.1007/s10997-013-9272-7>
- Soleimanof, S., Rutherford, M. W., & Webb, J. W. (2018). The Intersection of Family Firms and Institutional Contexts: A Review and Agenda for Future Research [Publisher: SAGE Publications Inc]. *Family Business Review*, 31(1), 32–53. <https://doi.org/10.1177/0894486517736446>
- Su, J., Zhai, Q., & Karlsson, T. (2017). Beyond Red Tape and Fools: Institutional Theory in Entrepreneurship Research, 1992–2014. *Entrepreneurship Theory and Practice*, 41(4), 505–531. Retrieved January 17, 2025, from <https://ideas.repec.org/a/sae/entthe/v41y2017i4p505-531.html>
- Sun, X., Lee, S.-H., & Phan, P. H. (2019). Family firm R&d investments in the 2007–2009 Great Recession. *Journal of Family Business Strategy*, 10. <https://doi.org/10.1016/J.JFBS.2018.02.004>
- Tokarczyk, J., Hansen, E., Green, M., & Down, J. (2007). A Resource-Based View and Market Orientation Theory Examination of the Role of “Familliness” in Family Business Success [Publisher: SAGE Publications Inc]. *Family Business Review*, 20(1), 17–31. <https://doi.org/10.1111/j.1741-6248.2007.00081.x>

- Ulland, A. M. L., & Upsaker, T. (2020). *Performance of Nordic public family firms : Comparable analysis of Nordic public family firm's performance against the Fama and french five-factor model* [Master's thesis] [Accepted: 2021-04-07T10:34:11Z]. Retrieved December 23, 2024, from <https://openaccess.nhh.no/nhh-xmlui/handle/11250/2736571>
- Urbinati, A., Chiaroni, D., & Chiesa, V. (2017). Towards a new taxonomy of circular economy business models. *Journal of Cleaner Production*, *168*, 487–498. <https://doi.org/10.1016/j.jclepro.2017.09.047>
- van Aaken, D., Göbel, M., & Meindl, D. (2020). Monitor or Advise? How Family Involvement Affects Supervisory Board Roles in Family Firms. *Schmalenbach Business Review*, *72*(2), 193–224. <https://doi.org/10.1007/s41464-020-00088-w>
- Veldman, J. (2019). Boards and Sustainable Value Creation: The Legal Entity, Co-Determination and Other Means. *European Business Law Review*, *30*(Issue 2), 279–300. <https://doi.org/10.54648/EULR2019013>
- Villalonga, B., & Amit, R. (2006). How do family ownership, control and management affect firm value? *Journal of Financial Economics*, *80*(2), 385–417. <https://doi.org/10.1016/j.jfineco.2004.12.005>
- Villalonga, B., Amit, R., Trujillo, M.-A., & Guzmán, A. (2015). Governance of Family Firms. *Review of Financial Economics*, *7*, 635–654. <https://doi.org/10.1146/ANNUREV-FINANCIAL-110613-034357>
- Villalonga, B., & Amit, R. (2009). How Are U.S. Family Firms Controlled? *The Review of Financial Studies*, *22*(8), 3047–3091. <https://doi.org/10.1093/rfs/hhn080>
- Yeh, Y.-H., & Woidtke, T. (2005). Commitment or entrenchment?: Controlling shareholders and board composition. *Journal of Banking & Finance*, *29*(7), 1857–1885. <https://doi.org/10.1016/j.jbankfin.2004.07.004>
- Zhao, M., Tan, J., & Park, S. H. (2014). From Voids to Sophistication: Institutional Environment and MNC CSR Crisis in Emerging Markets. *Journal of Business Ethics*, *122*(4), 655–674. <https://doi.org/10.1007/s10551-013-1751-x>

8 Appendices

8.1 Appendix 1: Data Extraction

8.1.1 Appendix 1A: Variables

Table 11: Overview of Extracted Variables

Variable	Explanation
company	The official name of the company as registered, used to match with other datasets and ensure correct identification.
ticker	The stock ticker symbol under which the company is publicly traded, facilitating financial data collection from market sources.
SH - Name	The name of each shareholder, used to identify ownership parties and assess governance structures.
SH - Type	The type/category of the shareholder (e.g., individual, family, institutional), helping distinguish ownership groups.
SH - Direct %	The direct ownership percentage held by the shareholder, indicating their immediate control.
SH - Total %	The total ownership percentage, combining direct and indirect stakes, reflecting overall influence.
DM Full Name	The full name of each board member or director, used to assess governance composition.
DM Job Title (in English)	The job title of each board member or director (in English), providing role-specific governance details.
date_incorporated	The year the company was legally established, used to calculate firm age.
country_iso	The ISO-standardized country code indicating the company's primary location, allowing for regional classification.
sic_3_digit	The three-digit Standard Industrial Classification (SIC) code categorizing the company's main industry activity.
market_cap_usd	The total market value of the company's outstanding shares, measured in US dollars, reflecting firm size and investor valuation.
total_assets_th	The book value of all company assets, reported in thousands, indicating the scale of operations.
sales_th_2024	The company's net sales (revenue) for the fiscal year 2024, reported in thousands, used to track recent performance.
sales_th_2021	The company's net sales for the fiscal year 2021, providing historical sales data for growth analysis.
ebitda_th	Earnings Before Interest, Taxes, Depreciation, and Amortization (EBITDA), reported in thousands, reflecting core profitability.
liabilities_th	The total liabilities of the company, in thousands, indicating the scale of financial obligations.
long_term_debt_th	The portion of liabilities classified as long-term debt, in thousands, measuring long-term leverage.
short_term_debt_th	The portion of liabilities classified as short-term debt, in thousands, indicating near-term financial commitments.

(continued on next page)

Variable	Explanation
shares_outstanding_th	The total number of shares currently outstanding, in thousands, used to compute per-share financial ratios.
book_value_per_share_usd	The book value of equity per share, in US dollars, reflecting the accounting value of shareholder equity.
market_price_usd	The current market price per share, in US dollars, reflecting investor valuation in the stock market.
total_assets_th_prev_year	Total assets from the previous fiscal year, in thousands, used for calculating asset-based growth rates.

8.1.2 Appendix 1B: Method

The dataset was initially downloaded in multiple subsets due to export restrictions in the Orbis database, after which these subsets were combined into a single dataset. Duplicate entries were identified and removed to ensure data integrity. To organize the data, dictionaries were created for all shareholders, capturing their names, types, and ownership percentages, as well as for all board members, capturing their names and roles. Additionally, all shareholder records with less than 1% ownership were excluded from the dataset to focus the analysis on meaningful ownership stakes.

8.1.3 Appendix 1C: Filters Applied

Table 12: Overview of Orbis Filters Applied

Filter	Explanation
Active companies / Unknown situation	Only companies that are currently active or whose operational status is unknown were included, excluding dissolved or inactive firms.
Location: EU or USA	The sample was restricted to companies located in European Union countries or the United States, ensuring regional focus.
Publicly traded	Only publicly listed firms were selected, allowing for consistent access to financial and governance data.

8.2 Appendix 2: Countries

Table 13: Countries Included in the Sample

Country Code	Country Name
Europe	
AT	Austria
BE	Belgium
BG	Bulgaria
HR	Croatia
CY	Cyprus
CZ	Czech Republic
DK	Denmark
EE	Estonia
FI	Finland
FR	France
DE	Germany
GR	Greece
HU	Hungary
IE	Ireland
IT	Italy
LV	Latvia
LT	Lithuania
LU	Luxembourg
MT	Malta
NL	Netherlands
PL	Poland
PT	Portugal
RO	Romania
SK	Slovakia
SI	Slovenia
ES	Spain
SE	Sweden
United States	
US	United States

8.3 Appendix 3: Industry Categories

Table 14: SIC Division Structure

Division	SIC Range	Industry Description
A	01-09	Agriculture, Forestry, and Fishing
B	10-14	Mining
C	15-17	Construction
D	20-39	Manufacturing
E	40-49	Transportation and Public Utilities
F	50-51	Wholesale Trade
G	52-59	Retail Trade
H	60-67	Finance, Insurance, and Real Estate
I	70-89	Services
J	91-99	Public Administration
K	99	Nonclassifiable Establishments

8.4 Appendix 4: Financial Calculations

Table 15: Overview of Financial Metric Calculations

Metric	Formula	Description
Book Value of Total Assets	$\text{total_assets_th} * 1000$	Total assets (in thousands) scaled up to full value.
Book Value of Equity	$(\text{total_assets_th} * 1000) - (\text{liabilities_th} * 1000)$	Total assets minus total liabilities.
Market Value of Equity	$\text{shares_outstanding_th} * 1000 * \text{market_price_usd}$	Number of shares times market price per share.
Total Debt	$(\text{long_term_debt_th} + \text{short_term_debt_th}) * 1000$	Combined long-term and short-term debt, scaled.
Market Value of Total Assets	$\text{book_value_total_assets} - \text{book_value_equity} + \text{market_value_equity}$	Adjusted total assets using market values.
Tobin's Q	$(\text{market_value_equity} - \text{total_debt}) / \text{book_value_total_assets}$	Measures market-to-book asset value.
Average Book Value of Total Assets	$(\text{total_assets_th_prev_year} * 1000 + \text{book_value_total_assets}) / 2$	Average of current and prior year's total assets.
Return on Assets (ROA)	$(\text{ebitda_th} * 1000) / \text{average_book_value_total_assets}$	EBITDA divided by average total assets.
Industry Code (2-digit)	$\text{str}(\text{sic_3_digit})[:2]$	First two digits of SIC industry code.
Firm Age	$2024 - \text{year}(\text{date_incorporated})$	Years since incorporation.
Firm Size	$\log(\text{book_value_total_assets})$	Natural log of total assets.
Sales Growth (3 years)	$(\text{sales_th_2024} - \text{sales_th_2021}) / \text{sales_th_2021}$	Percentage change in sales over 3 years.
Leverage	$\text{total_debt} / \text{book_value_equity}$	Debt-to-equity ratio.

8.5 Appendix 5: Ruleset

Table 16: Overview of Family Firm Classification Ruleset (ruleset.b)

Filter	Description
F1: Ownership Threshold	Firm has at least one named individual or family shareholder holding more than 20% ownership, either direct or total.
F2: Shareholder–Board Member Name Matching	There is a fuzzy string match (similarity 80%) between shareholder names and board member names, indicating family presence in governance.
F3: Family-Related Keywords	Shareholder names contain family-related keywords such as “Family,” “Holding,” “Trust,” “Foundation,” “Group,” “Partners,” “Invest,” “Enterprises,” “Holdings,” or “Inheritance.”
Final Rule	Firm is classified as a family firm if it satisfies F1 **and** (F2 **or** F3), ensuring that both ownership control and family presence are captured.

8.6 Appendix 6: Frequencies

Table 17: Descriptive Summary of Sample by Industry and Family Status (Ordered by SIC Code)

Industry Group	Family Firms	Non-family Firms	Total
Europe			
Construction	9	50	59
Manufacturing	117	556	673
Mining	3	27	30
Public Administration	2	11	13
Retail Trade	17	57	74
Services	73	295	368
Transportation and Public Utilities	29	161	190
Wholesale Trade	25	71	96
Total Europe	275	1228	1503
United States			
Construction	10	42	52
Manufacturing	109	968	1077
Mining	13	112	125
Nonclassifiable Establishments	0	2	2
Public Administration	0	2	2
Retail Trade	31	159	190
Services	138	638	776
Transportation and Public Utilities	25	185	210
Wholesale Trade	16	73	89
Total United States	342	2179	2521

8.7 Appendix 7: Codes

8.7.1 Appendix 7A; Dataset Assembly

```
1 import pandas as pd
2 from rapidfuzz import fuzz
3 import glob
4 import os
5
6 # === Step 1: Load and combine all Excel files ===
7 file_paths = glob.glob(r"C:\Users\mauri\Downloads\final_raw_dataset_*.xlsx")
8
9 if not file_paths:
10     raise FileNotFoundError("No files matched the pattern. Check your path and filenames.
11     ")
12 print(f" Found {len(file_paths)} files:")
13 for path in file_paths:
14     print("          ", os.path.basename(path))
15
16 combined_df_list = []
```

```

17 for file in file_paths:
18     df_temp = pd.read_excel(file, sheet_name='Results')
19     print(f" Loaded {len(df_temp)} rows from: {os.path.basename(file)}")
20     df_temp['Source File'] = os.path.basename(file) # Optional: keep track of origin
21     combined_df_list.append(df_temp)
22
23 # Combine
24 df = pd.concat(combined_df_list, ignore_index=True)
25 print(f" Total combined rows before deduplication: {len(df)}")
26
27 # === Step 2: Cleaning ===
28 df[['Company name Latin alphabet', 'Ticker symbol']] = df[['Company name Latin alphabet',
29     'Ticker symbol']].ffill()
30 df = df[~(df['SH - Name'].isna() & df['DMFull name'].isna())]
31 df['SH - Direct %'] = pd.to_numeric(df['SH - Direct %'], errors='coerce')
32 df['SH - Total %'] = pd.to_numeric(df['SH - Total %'], errors='coerce')
33
34 pre_filter_len = len(df)
35
36 # Remove shareholders with both Direct and Total ownership below 1%
37 df = df[
38     (df['SH - Direct %'].fillna(0) >= 1) |
39     (df['SH - Total %'].fillna(0) >= 1) |
40     (df['DMFull name'].notna()) # Keep rows with valid decision makers
41 ]
42 print(f" Rows after filtering shareholders with <1% ownership: {len(df)}")
43
44 # === Step 3: Build shareholder + DM strings ===
45 df['Shareholder Info'] = df.apply(
46     lambda row: f"{row['SH - Name']} (Type: {row['SH - Type']}, Direct: {row['SH - Direct
47     %']}, Total: {row['SH - Total %']})"
48     if pd.notna(row['SH - Name']) else None,
49     axis=1
50 )
51 df['DM Info'] = df.apply(
52     lambda row: f"{row['DMFull name']} - {row['DMJob title (in English)]}"
53     if pd.notna(row['DMFull name']) else None,
54     axis=1
55 )
56
57 # === Step 4: Rename financial variables ===
58 rename_cols = {
59     'Date of incorporation': 'date_incorporated',
60     'Country ISO code': 'country_iso',
61     'US SIC, core code (3 digits)': 'sic_3_digit',
62     'Market capitalisation\nm USD Last avail. yr': 'market_cap_usd',
63     'Shares outstanding\nth Last avail. yr': 'shares_outstanding_th',
64     'Book Value per share\nUSD Last avail. yr': 'book_value_per_share_usd',
65     "Shareholders funds\nth USD Last avail. yr": "shareholders_funds_th",
66     "Total assets\nth USD Last avail. yr": "total_assets_th",
67     "Long term debt\nth USD Last avail. yr": "long_term_debt_th",
68     "Loans & short-term debt\nth USD Last avail. yr": "short_term_debt_th",
69     "EBITDA\nth USD Last avail. yr": "ebitda_th",
70     "Research & Development expenses\nth USD Last avail. yr": "r_and_d_expenses_th",
71     "Sales\nth USD 2024": "sales_th_2024",
72     "Sales\nth USD 2023": "sales_th_2023",
73     "Sales\nth USD 2022": "sales_th_2022",
74     "Sales\nth USD 2021": "sales_th_2021",
75     "Total shareholders' equity\nth USD Last avail. yr": "equity_th",

```

```

74     "Total shareholders' equity\nth USD 2024": "equity_th_2024",
75     "Total shareholders' equity\nth USD 2023": "equity_th_2023",
76     "Total shareholders' equity\nth USD 2022": "equity_th_2022",
77     "Total shareholders' equity\nth USD 2021": "equity_th_2021",
78     "Total liabilities\nth USD Last avail. yr": "liabilities_th",
79     "Total liabilities\nth USD 2024": "liabilities_th_2024",
80     "Total liabilities\nth USD 2023": "liabilities_th_2023",
81     "Total liabilities\nth USD 2022": "liabilities_th_2022",
82     "Total liabilities\nth USD 2021": "liabilities_th_2021",
83     "Market price - year end\nUSD Last avail. yr": "market_price_usd",
84     "Operating profit (loss) [EBIT]\nth USD Last avail. yr": "ebit_th",
85     "Operating profit (loss) [EBIT]\nth USD 2024": "ebit_th_2024",
86     "Total assets\nth USD Year - 1": "total_assets_th_prev_year"
87 }
88 df.rename(columns=rename_cols, inplace=True)
89
90 # === Step 5: Group to company level and retain financials ===
91 financial_cols = list(rename_cols.values())
92
93 financial_info = df.groupby(['Company name Latin alphabet', 'Ticker symbol'])[
94     financial_cols].first().reset_index()
95 governance_info = df.groupby(['Company name Latin alphabet', 'Ticker symbol']).agg({
96     'Shareholder Info': lambda x: [i for i in x.dropna()],
97     'DM Info': lambda x: [i for i in x.dropna()]
98 }).reset_index()
99 # === Step 6: Merge and apply filters ===
100 grouped_df = pd.merge(governance_info, financial_info, on=['Company name Latin alphabet',
101     'Ticker symbol'], how='left')
102 grouped_df.rename(columns={
103     'Company name Latin alphabet': 'company',
104     'Ticker symbol': 'ticker',
105     'Shareholder Info': 'shareholders',
106     'DM Info': 'board_members',
107 }, inplace=True)
108 # Drop duplicates
109 df = grouped_df.drop_duplicates(subset=['company', 'ticker'])
110 print(f" Rows after drop_duplicates(): {len(df)}")
111 print(f" Rows dropped as duplicates: {len(grouped_df) - len(df)}")
112
113 # === Step 7: Reorder columns: General -> Shareholder -> Board -> Financials ===
114 final_column_order = [
115     'company',                # General
116     'ticker',
117     'date_incorporated',
118     'country_iso',
119     'sic_3_digit',
120     'shareholders',          # Shareholder
121     'board_members',         # Board
122     # Financials
123     'market_cap_usd',
124     'shares_outstanding_th',
125     'book_value_per_share_usd',
126     'shareholders_funds_th',
127     'total_assets_th',
128     'long_term_debt_th',
129     'short_term_debt_th',
130     'ebitda_th',

```

```

131     'r_and_d_expenses_th',
132     'sales_th_2024',
133     'sales_th_2023',
134     'sales_th_2022',
135     'sales_th_2021',
136     'equity_th',
137     'equity_th_2024',
138     'equity_th_2023',
139     'equity_th_2022',
140     'equity_th_2021',
141     'liabilities_th',
142     'liabilities_th_2024',
143     'liabilities_th_2023',
144     'liabilities_th_2022',
145     'liabilities_th_2021',
146     'market_price_usd',
147     'ebit_th',
148     'ebit_th_2024',
149     'total_assets_th_prev_year'
150 ]
151
152 # Ensure all columns exist before reordering (handles any missing values gracefully)
153 existing_cols = [col for col in final_column_order if col in df.columns]
154 df = df[existing_cols]
155
156 print(" Final DataFrame columns ordered by: General > Shareholder > Board > Financials")
157
158 # === Step 8: Export to Excel ===
159 output_path = r"C:\Users\mauri\Downloads\final_dataset.xlsx"
160 df.to_excel(output_path, index=False)
161 print(f" File successfully saved to: {output_path}")

```

Listing 1: Dataset Assembly

8.7.2 Appendix 7B: Datasample Assembly

```

1 import pandas as pd
2
3 # === Step 1: Load the final dataset ===
4 df = pd.read_excel(r"C:\Users\mauri\Downloads\final_dataset.xlsx")
5
6 print(f" Loaded dataset with {len(df)} rows.")
7
8 # === Step 2: Exclude Financial Institutions (SIC 600 699 ) ===
9 df['sic_3_digit'] = pd.to_numeric(df['sic_3_digit'], errors='coerce')
10 pre_filter = len(df)
11 df = df[~df['sic_3_digit'].between(600, 699)]
12 print(f" Excluded financial institutions: {pre_filter - len(df)} rows dropped")
13
14 # === Step 3: Exclude Countries ===
15
16 # Define countries to exclude
17 exclude_countries = ['CH', 'GB'] # CH = Switzerland, GB = United Kingdom
18
19 # Drop all non-EU members, except for the USA
20 eu_members = [
21     'AT', 'BE', 'BG', 'HR', 'CY', 'CZ', 'DK', 'EE', 'FI', 'FR', 'DE', 'GR', 'HU', 'IE', '
    IT', 'LV',

```

```

22     'LT', 'LU', 'MT', 'NL', 'PL', 'PT', 'RO', 'SK', 'SI', 'ES', 'SE'
23 ]
24
25 df = df[df['country_iso'].isin(eu_members + ['US'])]
26 print(f" After filtering by country: {len(df)} rows remain")
27
28 # === Step 4: Exclude companies with missing critical financials ===
29
30 # Convert "n.a", "N/A", etc. to actual NaN values
31 df.replace(to_replace=['n.a.', 'N/A', 'NA', 'na', 'NaN', 'nan'], value=pd.NA, inplace=
    True)
32
33 df['sales_th_2021'] = df['sales_th_2021'].replace(0, pd.NA)
34
35 # Decide which financial columns are essential
36 critical_financials = [
37     'market_cap_usd', 'total_assets_th', 'sales_th_2024', 'sales_th_2021', 'ebitda_th',
38     'liabilities_th', 'long_term_debt_th', 'short_term_debt_th', 'shares_outstanding_th',
39     'book_value_per_share_usd', 'market_price_usd', 'total_assets_th_prev_year',
40 ]
41
42 missing_counts = df[critical_financials].isnull().sum()
43 print(" Missing values in essential financials:")
44 print(missing_counts)
45
46 # Drop rows with missing essential data
47 pre_drop = len(df)
48 df = df.dropna(subset=critical_financials)
49 print(f" Companies dropped due to missing financials: {pre_drop - len(df)}")
50
51 # Re-convert financial columns to numeric after replacement
52 numeric_cols = [
53     'market_cap_usd', 'total_assets_th', 'sales_th_2024', 'sales_th_2021', 'ebitda_th',
54     'liabilities_th', 'long_term_debt_th', 'short_term_debt_th', 'shares_outstanding_th',
55     'book_value_per_share_usd', 'market_price_usd', 'total_assets_th_prev_year'
56 ]
57 df[numeric_cols] = df[numeric_cols].apply(pd.to_numeric, errors='coerce')
58
59 # === Step: Keep only necessary columns ===
60
61 # Column mapping to internal dataset
62 final_columns = [
63     'company',           # name
64     'ticker',            # ticker
65     'shareholders',     # governance
66     'board_members',    # governance
67     'date_incorporated',
68     'country_iso',      # year of incorporation
69     'sic_3_digit',
70     # Financials
71     'market_cap_usd',
72     'total_assets_th',
73     'sales_th_2024',
74     'sales_th_2021',
75     'ebitda_th',
76     'liabilities_th',
77     'long_term_debt_th',
78     'short_term_debt_th',
79     'shares_outstanding_th',

```

```

80     'book_value_per_share_usd',
81     'market_price_usd',
82     'total_assets_th_prev_year',
83     'r_and_d_expenses_th'
84 ]
85
86 # Keep only columns that exist
87 existing_final_columns = [col for col in final_columns if col in df.columns]
88 df = df[existing_final_columns]
89
90 print(f" Final dataset includes {len(df.columns)} columns and {len(df)} rows.")
91
92 # === Step 5: Export cleaned dataset ===
93 output_path = r"C:\Users\mauri\Downloads\final_sample.xlsx"
94 df.to_excel(output_path, index=False)
95 print(f" Final filtered file saved to: {output_path}")

```

Listing 2: Datasample Assembly

8.7.3 Appendix 7C: Datasample Enhanced

```

1 import pandas as pd
2 import pandas as pd
3 import numpy as np
4 from rapidfuzz import fuzz
5
6 # === Step 1: Load the final dataset ===
7 df = pd.read_excel(r"C:\Users\mauri\Downloads\final_sample.xlsx")
8
9 print(f" Loaded dataset with {len(df)} rows.")
10
11 # === Define EU members and assign region ===
12 eu_members = [
13     'AT', 'BE', 'BG', 'HR', 'CY', 'CZ', 'DK', 'EE', 'FI', 'FR', 'DE', 'GR', 'HU',
14     'IE', 'IT', 'LV', 'LT', 'LU', 'MT', 'NL', 'PL', 'PT', 'RO', 'SK', 'SI', 'ES', 'SE'
15 ]
16
17 # Add region column
18 df['region'] = df['country_iso'].apply(lambda x: 'USA' if x == 'US' else ('EU' if x in
19     eu_members else 'Other'))
20
21 # === Optional: Reorder to place 'region' right after 'country_iso' ===
22 cols = list(df.columns)
23 if 'country_iso' in cols and 'region' in cols:
24     idx = cols.index('country_iso') + 1
25     # Remove 'region' and reinsert at desired location
26     cols.insert(idx, cols.pop(cols.index('region')))
27     df = df[cols]
28
29 print(" Added 'region' column (USA / EU) based on country_iso.")
30
31 # Financial metric calculations ===
32 df["book_value_total_assets"] = df["total_assets_th"] * 1000
33
34 df["book_value_equity"] = df["total_assets_th"] * 1000 - df["liabilities_th"] * 1000
35
36 df["market_value_equity"] = df["shares_outstanding_th"] * 1000 * df["market_price_usd"]

```

```

37 df["total_debt"] = ( df["long_term_debt_th"] + df["short_term_debt_th"] ) * 1000
38
39 # Market Value of Total Assets
40 df["market_value_total_assets"] = df["book_value_total_assets"] - df["book_value_equity"]
    + df["market_value_equity"]
41
42 # Tobin's Q
43 df["tobins_q"] = ( df["market_value_equity"] - df["total_debt"] ) / df["
    book_value_total_assets"]
44
45 df["average_book_value_total_assets"] = ( df["total_assets_th_prev_year"] * 1000 + df["
    book_value_total_assets"] ) / 2
46
47 # ROA (Return on Assets)
48 df["roa"] = df["ebitda_th"] * 1000 / df["average_book_value_total_assets"]
49
50 # Industry code 2-digit
51 df["industry_code_2digit"] = df["sic_3_digit"].astype(str).str[:2]
52
53 # Firm Age
54 df["firm_age"] = 2024 - pd.to_datetime(df["date_incorporated"], errors='coerce').dt.year
55 df = df.dropna(subset=['firm_age'])
56
57 # Firm Size (log of assets)
58 df["firm_size"] = np.log(df["book_value_total_assets"])
59
60 # Growth in sales over 3 years
61 df["growth_3y_sales"] = (df["sales_th_2024"] - df["sales_th_2021"]) / df["sales_th_2021"]
62
63 # Leverage
64 df["leverage"] = df["total_debt"] / df["book_value_equity"]
65
66 # === Step 5: Export cleaned dataset ===
67 output_path = r"C:\Users\mauri\Downloads\final_sample_enhanced.xlsx"
68 df.to_excel(output_path, index=False)
69 print(f" Final filtered file saved to: {output_path}")

```

Listing 3: Dataset Enhanced

8.7.4 Appendix 7D: Datasample Categorized

```

1 import pandas as pd
2 import numpy as np
3 from rapidfuzz import fuzz
4 import ast
5 from scipy.stats.mstats import winsorize
6
7 # === Load your dataset ===
8 df = pd.read_excel(r"C:\Users\mauri\Downloads\final_sample_enhanced.xlsx")
9
10 print(f" Loaded dataset with {len(df)} rows.")
11
12 # Fix for stringified lists
13 def safe_parse_list(val):
14     try:
15         return ast.literal_eval(val) if isinstance(val, str) else val
16     except:
17         return []

```

```

18
19 df['shareholders'] = df['shareholders'].apply(safe_parse_list)
20 df['board_members'] = df['board_members'].apply(safe_parse_list)
21
22
23 # === Define family keywords ===
24 FAMILY_KEYWORDS = {
25     "FAMILY", "HOLDING", "TRUST", "FOUNDATION", "GROUP",
26     "PARTNERS", "INVEST", "ENTERPRISES", "HOLDINGS", "INHERITANCE"
27 }
28
29 # === Helper: Extract % values from string ===
30 def extract_percent(text, key):
31     try:
32         if f"{key}:" in text:
33             val = text.split(f"{key}:")[1].split(",")[0].replace("%", "").strip()
34             return float(val) if val.lower() != "nan" else None
35     except:
36         return None
37     return None
38
39 # === Filter 1: Ownership > 20% and individual/family ===
40 def check_f1(shareholders):
41     if not isinstance(shareholders, list):
42         return 0
43     for sh in shareholders:
44         if "One or more named individuals or families" in sh:
45             direct = extract_percent(sh, "Direct")
46             total = extract_percent(sh, "Total")
47             if (direct is not None and direct > 20) or (total is not None and total > 20)
48                 :
49                 return 1
50     return 0
51 df['F1'] = df['shareholders'].apply(check_f1)
52
53 # === Fuzzy name match: shareholder ~ board member ===
54 def fuzzy_match(sh_list, dm_list, threshold=80):
55     if not isinstance(sh_list, list) or not isinstance(dm_list, list):
56         return 0
57     for sh in sh_list:
58         sh_name = sh.split(" (")[0].strip().lower()
59         for dm in dm_list:
60             dm_name = dm.split(" -")[0].strip().lower()
61             if fuzz.token_sort_ratio(sh_name, dm_name) >= threshold:
62                 return 1
63     return 0
64
65 df['F2'] = df.apply(lambda row: fuzzy_match(row['shareholders'], row['board_members']),
66                    axis=1)
67
68 # === Filter 3: shareholder name contains family keyword ===
69 def contains_keyword(name_list):
70     if not isinstance(name_list, list):
71         return 0
72     return int(any(any(kw in name.upper() for kw in FAMILY_KEYWORDS) for name in
73                  name_list))
74 df['F3'] = df['shareholders'].apply(contains_keyword)

```

```

74
75 def check_f4(shareholders):
76     if not isinstance(shareholders, list):
77         return 0
78     for sh in shareholders:
79         if "One or more named individuals or families" in sh:
80             direct = extract_percent(sh, "Direct")
81             total = extract_percent(sh, "Total")
82             name = sh.split(" ")[0].strip().upper()
83             if ((direct and direct > 20) or (total and total > 20)) and any(kw in name
for kw in FAMILY_KEYWORDS):
84                 return 1
85     return 0
86
87 df['F4'] = df['shareholders'].apply(check_f4)
88
89 def check_f5(shareholders, board_members):
90     if not isinstance(shareholders, list) or not isinstance(board_members, list):
91         return 0
92     for sh in shareholders:
93         if "One or more named individuals or families" in sh:
94             direct = extract_percent(sh, "Direct")
95             total = extract_percent(sh, "Total")
96             sh_name = sh.split(" ")[0].strip().lower()
97             if (direct and direct > 20) or (total and total > 20):
98                 for dm in board_members:
99                     dm_name = dm.split(" ")[0].strip().lower()
100                    if fuzz.token_sort_ratio(sh_name, dm_name) >= 80: # Higher match
threshold for precision
101                        return 1
102     return 0
103
104 df['F5'] = df.apply(lambda row: check_f5(row['shareholders'], row['board_members']), axis
=1)
105
106 def check_f6(shareholders, board_members):
107     if not isinstance(shareholders, list) or not isinstance(board_members, list):
108         return 0
109     for sh in shareholders:
110         sh_name = sh.split(" ")[0].strip()
111         if any(kw in sh_name.upper() for kw in FAMILY_KEYWORDS):
112             for dm in board_members:
113                 dm_name = dm.split(" ")[0].strip().lower()
114                 if fuzz.token_sort_ratio(sh_name.lower(), dm_name) >= 80:
115                     return 1
116     return 0
117
118 df['F6'] = df.apply(lambda row: check_f6(row['shareholders'], row['board_members']), axis
=1)
119
120 def check_f7(shareholders, board_members):
121     if not isinstance(shareholders, list) or not isinstance(board_members, list):
122         return 0
123     for sh in shareholders:
124         if "One or more named individuals or families" in sh:
125             direct = extract_percent(sh, "Direct")
126             total = extract_percent(sh, "Total")
127             name = sh.split(" ")[0].strip()
128             name_upper = name.upper()

```

```

129         if (direct and direct > 20) or (total and total > 20):
130             if any(kw in name_upper for kw in FAMILY_KEYWORDS):
131                 for dm in board_members:
132                     dm_name = dm.split(" -")[0].strip().lower()
133                     if fuzz.token_sort_ratio(name.lower(), dm_name) >= 85:
134                         return 1
135         return 0
136
137 df['F7'] = df.apply(lambda row: check_f7(row['shareholders'], row['board_members']), axis
138                    =1)
139
140 # === RuleSets ===
141 # RuleSets
142 df['ruleset_a'] = ((df['F1'] == 1) | (df['F2'] == 1)).astype(int)
143 df['ruleset_b'] = ((df['F1'] == 1) & ((df['F2'] == 1) | (df['F3'] == 1))).astype(int)
144 df['ruleSet_c'] = ((df['F4'] == 1) | (df['F5'] == 1) | (df['F6'] == 1)).astype(int)
145 df['ruleSet_d'] = (df['F7']).astype(int)
146
147 # Family scores
148 df['family_score'] = 0
149 df.loc[df['ruleset_a'] == 1, 'family_score'] = 1
150 df.loc[df['ruleset_b'] == 1, 'family_score'] = 2
151 df.loc[df['ruleSet_c'] == 1, 'family_score'] = 3
152 df.loc[df['ruleSet_d'] == 1, 'family_score'] = 4
153
154 winsor_vars = ['tobins_q', 'roa', 'growth_3y_sales', 'leverage']
155 for var in winsor_vars:
156     if var in df.columns:
157         clean_vals = df[var].dropna()
158         winsorized = winsorize(clean_vals, limits=[0.01, 0.01])
159         df.loc[clean_vals.index, f"{var}_wins"] = winsorized
160
161 # === Export ===
162 output_path = r"C:\Users\mauri\Downloads\vt_sample_categorized.xlsx"
163 df.to_excel(output_path, index=False)
164 print(f" Final classified file saved to: {output_path}")

```

Listing 4: Dataset Categorized

8.7.5 Appendix 7E: Descriptive Analysis

```

1 import pandas as pd
2 import numpy as np
3 from scipy.stats import ttest_ind
4 from scipy.stats.mstats import winsorize
5
6 # === Step 1: Load dataset ===
7 df = pd.read_excel(r"C:\Users\mauri\Downloads\vt_sample_categorized.xlsx")
8
9 # === Step 2: Family firm classification ===
10 df['is_family_firm'] = df['family_score'] == 2
11
12 # === Step 3: Define variables and winsorize where needed ===
13 all_vars = ['tobins_q_wins', 'roa_wins', 'firm_age', 'firm_size', 'growth_3y_sales_wins',
14            'leverage_wins']
15
16 # === Step 4: Summary functions ===

```

```

16 def stars(p):
17     if p < 0.01: return '***'
18     elif p < 0.05: return '**'
19     elif p < 0.1: return '*'
20     return ''
21
22 def describe_group(var, fam, nonfam):
23     tstat, pval = ttest_ind(fam, nonfam, equal_var=False, nan_policy='omit')
24     return {
25         'Variable': var,
26         'Family Mean': fam.mean(),
27         'Family Median': fam.median(),
28         'Family S.E.': fam.sem(),
29         'Family Min': fam.min(),
30         'Family Max': fam.max(),
31         'Family N': len(fam),
32         'Non-Family Mean': nonfam.mean(),
33         'Non-Family Median': nonfam.median(),
34         'Non-Family S.E.': nonfam.sem(),
35         'Non-Family Min': nonfam.min(),
36         'Non-Family Max': nonfam.max(),
37         'Non-Family N': len(nonfam),
38         'T-Statistic': round(tstat, 2),
39         'P-Value': f"{pval:.3g}{stars(pval)}"
40     }
41
42 # === Step 5: Global Summary ===
43 summary_rows = []
44 for var in all_vars:
45     fam = df[df['is_family_firm']][var].dropna()
46     nonfam = df[~df['is_family_firm']][var].dropna()
47     summary_rows.append(describe_group(var, fam, nonfam))
48
49 global_summary = pd.DataFrame(summary_rows)
50
51 # === Step 6: Region-specific Summary ===
52 region_summaries = []
53 for region in ['EU', 'USA']:
54     df_region = df[df['region'] == region]
55     for var in all_vars:
56         fam = df_region[df_region['is_family_firm']][var].dropna()
57         nonfam = df_region[~df_region['is_family_firm']][var].dropna()
58         row = describe_group(var, fam, nonfam)
59         row['Region'] = region
60         region_summaries.append(row)
61
62 region_summary = pd.DataFrame(region_summaries)
63
64 # === Step 7: Custom Descriptive Stats (Full Dataset) ===
65 descriptive_rows = []
66 for var in all_vars:
67     series = df[var].dropna()
68     row = {
69         'Variable': var,
70         'Mean': series.mean(),
71         'Median': series.median(),
72         'Min': series.min(),
73         'Max': series.max(),
74         'S.E.': series.sem()

```

```

75     }
76     descriptive_rows.append(row)
77
78 overall_descriptive = pd.DataFrame(descriptive_rows)
79
80 # === Step 8: Add Count Tables ===
81 family_counts = df['is_family_firm'].value_counts().rename(index={True: 'Family Firms',
82     False: 'Non-Family Firms'})
83 family_counts = family_counts.rename('Count').reset_index().rename(columns={'index': '
84     Firm Type'})
85
86 region_family_counts = (
87     df.groupby(['region', 'is_family_firm'])
88     .size()
89     .unstack(fill_value=0)
90     .rename(columns={True: 'Family Firms', False: 'Non-Family Firms'})
91     .reset_index()
92 )
93
94 # === Step 9: Export all to Excel ===
95 output_path = r"C:\Users\mauri\Downloads\vt_summary_statistics.xlsx"
96
97 with pd.ExcelWriter(output_path, engine='xlsxwriter') as writer:
98     global_summary.to_excel(writer, sheet_name='Overall Summary', index=False)
99     region_summary.to_excel(writer, sheet_name='Regional Summary', index=False)
100    family_counts.to_excel(writer, sheet_name='Family vs Non-Fam Count', index=False)
101    region_family_counts.to_excel(writer, sheet_name='Region Fam Count', index=False)
102    overall_descriptive.to_excel(writer, sheet_name='Full Dataset Descriptives', index=
103    False)
104
105 print(f" Export complete: {output_path}")

```

Listing 5: Descriptive Analysis

8.7.6 Appendix 7F: Linear Regression

```

1 import pandas as pd
2 import statsmodels.formula.api as smf
3 from statsmodels.stats.outliers_influence import variance_inflation_factor
4 import statsmodels.api as sm
5 from scipy.stats import shapiro
6 from statsmodels.stats.diagnostic import het_breuschpagan
7
8 # === Step 1: Load dataset ===
9 df = pd.read_excel(r"C:\Users\mauri\Downloads\vt_sample_categorized.xlsx")
10
11 # === Step 2: Add family firm dummy, region dummy, and map industry ===
12 df['is_family_firm'] = df['family_score'] == 2
13 df['region_USA'] = (df['region'] == 'USA').astype(int)
14
15 def map_sic_division(code):
16     if 1 <= code <= 9:
17         return 'Agriculture, Forestry, and Fishing'
18     elif 10 <= code <= 14:
19         return 'Mining'
20     elif 15 <= code <= 17:
21         return 'Construction'
22     elif 20 <= code <= 39:

```

```

23     return 'Manufacturing'
24 elif 40 <= code <= 49:
25     return 'Transportation and Public Utilities'
26 elif 50 <= code <= 51:
27     return 'Wholesale Trade'
28 elif 52 <= code <= 59:
29     return 'Retail Trade'
30 elif 60 <= code <= 67:
31     return 'Finance, Insurance, and Real Estate'
32 elif 70 <= code <= 89:
33     return 'Services'
34 elif 91 <= code <= 99:
35     return 'Public Administration'
36 else:
37     return 'Nonclassifiable Establishments'
38
39 df['industry_group'] = df['industry_code_2digit'].apply(map_sic_division)
40
41 # === Step 3: Define variables ===
42 dependent_vars = ['tobins_q_wins', 'roa_wins']
43 control_vars = ['firm_size', 'firm_age', 'growth_3y_sales_wins', 'leverage_wins']
44 vif_vars = ['is_family_firm'] + control_vars
45 industry_fe = 'C(industry_group)'
46
47 # === Step 4: Helper functions ===
48 def stars(p):
49     if p < 0.01: return '***'
50     elif p < 0.05: return '**'
51     elif p < 0.1: return '*'
52     return ''
53
54 def interpret_result(coef, pval, region, var):
55     sig = pval < 0.1
56     if not sig:
57         return f"No significant effect of family firms on {var.replace('_wins', '')} in {
region}"
58     direction = "higher" if coef > 0 else "lower"
59     return f"Family firms have {direction} {var.replace('_wins', '')} in {region} (p < {
pval:.2g})"
60
61 def run_model(data, y_var):
62     formula = f"{y_var} ~ is_family_firm + {' + '.join(control_vars)} + {industry_fe}"
63     model_data = data[[y_var, 'is_family_firm'] + control_vars + ['industry_group']].
dropna()
64     model = smf.ols(formula, data=model_data).fit(cov_type='HC3')
65     return model, model_data
66
67 def run_diagnostics(model, model_data, y_var):
68     residuals = model.resid
69     shapiro_stat, shapiro_p = shapiro(residuals)
70     lm_stat, lm_pvalue, _, _ = het_breuschpagan(residuals, model.model.exog)
71     return {
72         'Dependent Variable': y_var.replace('_wins', ''),
73         'Shapiro-Wilk Statistic': round(shapiro_stat, 4),
74         'Shapiro-Wilk p-value': round(shapiro_p, 4),
75         'Breusch-Pagan LM Statistic': round(lm_stat, 4),
76         'Breusch-Pagan p-value': round(lm_pvalue, 4)
77     }
78

```

```

79 def calculate_vif(df, features):
80     X = sm.add_constant(df[features].astype(float))
81     vif_data = pd.DataFrame()
82     vif_data["Variable"] = X.columns
83     vif_data["VIF"] = [variance_inflation_factor(X.values, i) for i in range(X.shape[1])]
84     return vif_data
85
86 # === Step 5: Run models and collect all results ===
87 all_results = []
88 family_effects = []
89 diagnostic_results = []
90 vif_results = []
91
92 for region in ['Full Sample', 'EU', 'USA']:
93     df_region = df if region == 'Full Sample' else df[df['region'] == region]
94
95     for y in dependent_vars:
96         model_ctrl, model_data_ctrl = run_model(df_region, y)
97         n_obs = int(model_ctrl.nobs)
98         r2_ctrl = model_ctrl.rsquared
99         adj_r2_ctrl = model_ctrl.rsquared_adj
100
101         for var in model_ctrl.params.index:
102             coef = model_ctrl.params[var]
103             pval = model_ctrl.pvalues[var]
104             tstat = model_ctrl.tvalues[var]
105             se = model_ctrl.bse[var]
106             significance = stars(pval)
107
108             all_results.append({
109                 'Region': region,
110                 'Model': 'With Controls',
111                 'Dependent Variable': y.replace('_wins', ''),
112                 'Variable': var,
113                 'Coefficient': round(coef, 4),
114                 'Std. Error': round(se, 4),
115                 'T-Statistic': round(tstat, 2),
116                 'P-Value': round(pval, 4),
117                 'Significance': significance,
118                 'R-squared': round(r2_ctrl, 4),
119                 'Adj R-squared': round(adj_r2_ctrl, 4),
120                 'Observations': n_obs
121             })
122
123         # Family firm effect summary
124         coef = model_ctrl.params['is_family_firm[T.True]']
125         pval = model_ctrl.pvalues['is_family_firm[T.True]']
126         tstat = model_ctrl.tvalues['is_family_firm[T.True]']
127         se = model_ctrl.bse['is_family_firm[T.True]']
128         significance = stars(pval)
129
130         family_effects.append({
131             'Region': region,
132             'Model': 'With Controls',
133             'Dependent Variable': y.replace('_wins', ''),
134             'Coefficient': round(coef, 4),
135             'Std. Error': round(se, 4),
136             'T-Statistic': round(tstat, 2),
137             'P-Value': round(pval, 4),

```

```

138         'Significance': significance,
139         'R-squared': round(r2_ctrl, 4),
140         'Adj R-squared': round(adj_r2_ctrl, 4),
141         'Interpretation': interpret_result(coef, pval, region, y),
142         'Observations': n_obs
143     })
144
145     # Diagnostics
146     diag = run_diagnostics(model_ctrl, model_data_ctrl, y)
147     diag['Region'] = region
148     diagnostic_results.append(diag)
149
150     # Clean VIFs
151     vif_data = calculate_vif(model_data_ctrl, vif_vars)
152     vif_data['Dependent Variable'] = y.replace('_wins', '')
153     vif_data['Region'] = region
154     vif_results.append(vif_data)
155
156     # === Step 6: Export to Excel ===
157     all_results_df = pd.DataFrame(all_results)
158     family_effects_df = pd.DataFrame(family_effects)
159     diagnostics_df = pd.DataFrame(diagnostic_results)
160     vif_df_full = pd.concat(vif_results, ignore_index=True)
161
162     with pd.ExcelWriter(r"C:\Users\mauri\Downloads\vt_thesis_regression_output.xlsx") as
163         writer:
164         all_results_df.to_excel(writer, sheet_name='Full Results', index=False)
165         family_effects_df.to_excel(writer, sheet_name='Family Firm Summary', index=False)
166         diagnostics_df.to_excel(writer, sheet_name='Model Diagnostics', index=False)
167         vif_df_full.to_excel(writer, sheet_name='VIF Scores', index=False)
168     print(" Regression outputs, diagnostics, and clean VIFs saved successfully.")

```

Listing 6: Linear Regression

8.7.7 Appendix 7G: Correlation Matrix

```

1  \label{appendix:7g_correlation_matrix}
2  import pandas as pd
3  import numpy as np
4  from scipy.stats import pearsonr
5
6  # === Step 1: Load dataset ===
7  df = pd.read_excel(r"C:\Users\mauri\Downloads\vt_sample_categorized.xlsx")
8
9  # === Step 2: Family firm classification ===
10 df['is_family_firm'] = df['family_score'] == 2
11
12 # === Step 3: Correlation Matrix with Significance ===
13 cols_to_corr = [
14     'is_family_firm', 'roa_wins', 'tobins_q_wins',
15     'firm_size', 'firm_age', 'growth_3y_sales', 'leverage'
16 ]
17
18 annot_matrix = pd.DataFrame(index=cols_to_corr, columns=cols_to_corr, dtype=str)
19
20 for col1 in cols_to_corr:
21     for col2 in cols_to_corr:

```

```

22     series1 = df[col1].dropna()
23     series2 = df[col2].dropna()
24     common_idx = series1.index.intersection(series2.index)
25
26     if len(common_idx) > 2:
27         r, p = pearsonr(series1.loc[common_idx], series2.loc[common_idx])
28         if p < 0.01:
29             annot = f"{r:.2f}***"
30         elif p < 0.05:
31             annot = f"{r:.2f}**"
32         elif p < 0.1:
33             annot = f"{r:.2f}*"
34         else:
35             annot = f"{r:.2f}"
36         annot_matrix.loc[col1, col2] = annot
37     else:
38         annot_matrix.loc[col1, col2] = ""
39
40 # === Step 4: Save Annotated Matrix ===
41 output_path_annot = r"C:\Users\mauri\Downloads\vt_correlation_matrix_with_stars_final.
42     xlsx"
43
44 annot_matrix.to_excel(output_path_annot)
45
46 print(f"\n Saved annotated correlations (with stars) to {output_path_annot}")

```

Listing 7: Correlation Matrix