


Female Underperformance Hypothesis Revisited: Methodological Review and Empirical Testing

SAGE Open
October-December 2023: 1–25
© The Author(s) 2023
DOI: 10.1177/21582440231218798
journals.sagepub.com/home/sgo


María José Ibáñez^{1,2} , Felipe Vásquez Lavin^{3,4}, and Roberto D. Ponce Oliva^{3,4} 

Abstract

Comparison between the performance of female and male-managed firms has long been a subject of research interest. Although the argument is that firms run by women have lower performance than those run by men, there is no agreement on the effects of managerial gender on companies' financial outcomes. This study conducts a methodological review of quantitative research on the relationship between female business leadership and firm performance from 2010 to 2020. This review identifies the most frequently used dependent and explanatory variables and econometric models in the literature. Most studies have not considered endogeneity bias in their model specifications; therefore, these results could be biased and unreliable. We select empirical models to test the female underperformance hypothesis using a sample of Chilean firms. Our findings suggest that managers' gender does not significantly affect business performance when endogeneity is addressed. Our methodological review reveals a significant gap in the research on female managers and firm performance in the Latin American context, and the empirical test provides new evidence in this vein.

Plain Language Summary

A Fresh Look at the Female Underperformance Hypothesis

Comparison between the performance of female and male-managed firms has long been a subject of research interest. Although the argument is that firms run by women have lower performance than those run by men, there is no agreement on the effects of managerial gender on companies' financial outcomes. This study conducts a methodological review of quantitative research on the relationship between female business leadership and firm performance from 2010 to 2020. This review identifies the most frequently used dependent and explanatory variables and econometric models in the literature. Most studies have not considered endogeneity bias in their model specifications; therefore, these results could be biased and unreliable. We select empirical models to test the female underperformance hypothesis using a sample of Chilean firms. Our findings suggest that managers' gender does not significantly affect business performance when endogeneity is addressed. Our methodological review reveals a significant gap in the research on female managers and firm performance in the Latin American context, and the empirical test provides new evidence in this vein.

Keywords

female underperformance hypothesis, female CEO, firm performance, methodological review

¹CENTRUM Católica Graduate Business School, Lima, Perú

²Pontificia Universidad Católica del Perú, Lima, Perú

³School of Business and Economics, Universidad del Desarrollo, Concepción, Chile

⁴Center of Applied Ecology and Sustainability (CAPES), Santiago, Chile

Corresponding Author:

María José Ibáñez, Pontificia Universidad Católica del Perú, Jirón Daniel Alomía Robles N° 125 - 129, Lima 15023, Perú.

Email: mjibanez@puccp.edu.pe



Introduction

Many studies have examined the relationship between women's presence in corporate leadership positions and firm performance (Gipson et al., 2017; Lam et al., 2013; Mohan, 2014). For over 30 years, researchers have compared the financial reports of companies owned by men and women (Cuba et al., 1983; Fischer, 1992; Johnson & Storey, 1993). This research field began in the area of entrepreneurship and rapidly extended to studying the performance of companies with women in management positions (Fairlie & Robb, 2009; Kolev, 2012). Over time, the claim that women-owned or -managed companies perform worse than those owned by men or with men as CEOs stuck in the literature (Du Rietz & Henrekson, 2000), and this assumption was labeled the Female Underperformance Hypothesis (FUH) (Dean et al., 2019).

Research on FUH has generated mixed results, with evidence supporting and rejecting this hypothesis (Bennouri et al., 2018; Martín-Ugedo et al., 2019; Singhathep & Pholphirul, 2015). Some studies have found that female management positively affects firm performance (Christiansen et al., 2016; Conyon & He, 2017; Moreno-Gómez et al., 2018; Noland et al., 2016), while others have found that female-managed companies perform worse (Lim et al., 2019; Singhathep & Pholphirul, 2015). For example, Menicucci et al. (2019) have found that women-managed hotels in Italy outperformed those managed by men in terms of hotel growth. Kaur and Singh (2019) have suggested that firms led by female CEOs are negatively related to firm performance. The third group of studies has indicated that managers' gender has no significant effect on firm performance (Singh et al., 2019; Unite et al., 2019; Vu et al., 2019). Kristanti and Iswandi (2019) have shown that gender diversity in leadership positions is insignificant for business performance. Considering that there isn't a clear agreement on the FUH in previous studies, we are motivated to take another look at this issue and gather new evidence. We test the FUH using state-of-the-art dependent and explanatory variables and econometric methods and apply them to new data (Brahma et al., 2021; Watson, 2020; Đặng et al., 2020).

The increasing number of women in business leadership positions in recent decades has fueled continued interest in evaluating the financial performance of women-led companies (Helfat et al., 2006; Noland et al., 2016). As many researchers have pointed out, accepting FUH reinforces stereotypical social constructions of gender, in which management positions are interpreted as being more compatible with a masculine image (Carmona et al., 2018; Cook & Goodman, 2006; Foley et al., 2005; Metcalfe, 2007). Most studies on FUH are quantitative and based on econometric methods using

only financial indicators as dependent variables (J. Chen et al., 2018; Isidro & Sobral, 2015; Moreno-Gómez et al., 2018). Research on this topic has been conducted using data from large publicly listed companies (Terjesen et al., 2009). This research methodology cannot encapsulate all dimensions of firm performance and its focus on large firms limits the findings to a small portion of the business ecosystem (INE, 2017, 2022; Krishnan & Park, 2005; Rodríguez-Domínguez et al., 2012). Given these limitations, we believe there is room for new studies on FUH that consider women's recent business participation and new business ecosystems.

This study conducts an extensive review of quantitative research methodologies that implement different estimations to test the FUH in female-managed firms. We identify the variables and econometric methods used in the literature and apply them to a sample of 2,323 Chilean firms. The firms in this study have many different characteristics, such as legal establishment, sector, size, and age, which reflect the heterogeneity of the Chilean business system. Our findings do not support FUH when the specification endogeneity bias is resolved. Under the correct specification of the econometric model, female-managed firms do not underperform compared with businesses run by men. We show that a methodological research design can induce imprecise conclusions if the endogeneity bias is not addressed or non-representative samples are used. The endogeneity problem has been recognized in some empirical studies, and the authors have explored techniques to address this issue. However, endogeneity bias is not the primary concern in most specifications to establish the effect of CEO's gender on firm performance. Although endogeneity bias may seem to be a non-critical issue in the exploration of FUH, we demonstrate that specification problems should be considered, at least to ensure that the results are robust under alternative estimation methods. We contribute to the literature on female business leadership by suggesting a novel research agenda with methodological recommendations for future studies.

Section 1 presents the theoretical framework of this study. Section 2 describes the methodological design of the study. Section 3 presents the results discussed in Section 4. Finally, the conclusions are presented in Section 5. Finally, the implications of our study are presented in Section 6.

Theoretical Framework

Researchers have compared the financial data of male- and female-led firms over 30 years (Cuba et al., 1983; Fischer, 1992; Johnson & Storey, 1993). This research began with entrepreneurship, comparing the performance of small firms initiated by both men and women

(Fairlie & Robb, 2009; Hisrich & Brush, 2019; Loscocco et al., 1991). This topic was later expanded to studying the financial performance of companies managed by women (Kolev, 2012). Recently, researchers have focused on measuring firm performance in companies in which women serve on the board of directors (Aggarwal et al., 2019; Rose, 2007). In this research on women and firm performance, an underlying assumption is that female-owned firms or firms managed by women have lower performance than those that are male-owned or where the CEO is male (Du Rietz & Henrekson, 2000). This assumption is known in the literature as the FUH (Dean et al., 2019; Watson, 2002; Westhead, 2003). The FUH has been studied in diverse contexts, including entrepreneurship, management, and corporate governance. However, the body of research in these three fields can be understood as literature on female leadership and its relationship with business and financial performance (Kotiranta et al., 2010). Most studies have not explicitly mentioned FUH as a motivation for research, although their findings either support or reject this hypothesis (Zolin et al., 2013).

The first studies addressing women and firm performance can be found in female entrepreneurship by the end of the 1970s and 1980s (Moore, 1990). As more women started businesses and began working independently, interest grew as to whether these women-run businesses performed as well as those created by men (Chaganti, 1986). Findings have shown that women-owned companies have lower performance than male-owned companies (C. G. Brush, 1992; Hisrich & Brush, 1987; Loscocco et al., 1991). Several studies with similar results have reinforced this idea. Currently, there is agreement that the types of business opportunities that women approach tend to be in industries with relatively low performance (e.g., retail and service) (Ahl, 2006; Ahl & Marlow, 2012; Gupta et al., 2019; Hundley, 2001; Sullivan & Meek, 2012).

Traditionally, a manager's position is associated with a masculine image, which is a persistent stereotype of company leadership (Ellemers & Nadal, 2018). Male-dominated organizations are often reluctant to accept women in leadership roles because of the lingering perception that women do not have the attributes necessary to be successful in roles and positions typically viewed as masculine (Grover, 2015; Heilman, 2012). The small number of women in corporate leadership positions supports these exclusionary tendencies (Nili, 2019). Some evidence has demonstrated that women have fewer opportunities to access managerial positions because of gender bias (Sarrió et al., 2002). The increased presence of women in management positions has motivated renewed interest in understanding the distinct characteristics of women managers and their ability to meet or

surpass the performance of their male counterparts (Helfat et al., 2006; Kanuri & Malm, 2018; Noland et al., 2016).

Studies of FUH have yielded contradictory findings. Many studies have shown that female leadership positively affects firm performance (Christiansen et al., 2016; Conyon & He, 2017; Moreno-Gómez et al., 2018; Noland et al., 2016), whereas others have maintained that women-run companies perform worse (Lim et al., 2019; Singhathep & Pholphirul, 2015). Others have claimed that a manager's gender has no significant effect on firm performance (Singh et al., 2019; Unite et al., 2019; Vu et al., 2019). Recently, some studies have explicitly challenged the FUH. Watson (2020) have indicated that the low performance of companies managed by women is a myth. Ho et al. (2015) have argued that female leaders' contributions to firm performance cannot be measured solely in economic terms. Academic interest in women who are successful in business leadership positions has grown slowly; thus, some studies have focused on these women and their abilities, instead of simply comparing the performance of different genders (C. Brush, 2019; Connell, 2019; Crittenden & Bliton, 2019).

In corporate governance, the representation of women in the board directory is of interest to policymakers. Some countries have introduced gender diversity requirements or quotas for boards of directors publicly traded companies (Terjesen et al., 2009). For example, in 2003, Norway adopted a gender quote requiring a 40% female board representation in publicly limited state-owned companies. In 2011, Italy implemented a directive that listed companies' boards should have at least 33% under-represented genders (ILO Bureau for Employers' Activities, 2020). Most studies on board gender diversity have related to firm performance; only a few have looked beyond financial results and included the demographic, human capital, and social capital contributions of women (Bear et al., 2010; Bennouri et al., 2018; Kirsch, 2018). Studying gender diversity in boards of directors involves many ethical considerations related to female participation in different social environments (Ntim, 2015; Porcena et al., 2021). Similar to female management, research on the gender diversity of directive boards and its impact on firm performance has not been conclusive (Reddy & Jadhav, 2019). For example, Post and Byron (2015; p.1546) have found that "female board representation is positively related to accounting returns and that this relationship is more positive in countries with stronger shareholder protections." Terjesen et al. (2016, p. 447) have suggested that "external independent directors do not contribute to firm performance unless the board is gender-diversified." Unite et al. (2019, p. 65) have found that female and male corporate leaders in Philippine firms have comparable competency levels and

that increasing the presence of women on corporate boards has no discernible effect on firm performance.

Considering that boards' decision-making processes include contributions from most board members (men and women), it is difficult to identify the specific contributions of female (or male) members and arrive at a deeper understanding of how that contribution influences firm performance (Strydom et al., 2017). Recently, Nekhili et al. (2018) have suggested that family firms have a significantly positive relationship between the appointment of a female chair and firm performance. In addition, a negative and significant relationship is observed between female chairs and return on assets in non-family firms. Gradually, this conversation has transitioned to researching the active role of women on boards, paying special attention to economies in which the continued dominance of patriarchy has delayed the participation of women at the highest levels of the corporate hierarchy (Farhan & Nayan, 2018; Green & Homroy, 2018).

Our research explores in depth the methodological issues applied to test the FUH in selected quantitative studies to answer the following research questions:

RQ1. What methods and variables have been used to test FUH?

RQ2. What conclusions do the results of previous quantitative studies offer regarding FUH?

RQ3. Do the findings and conclusions obtained from the FUH hold when alternative methods and data are used?

The Roots of FUH: Foundations and Evidence

FUH research has been conducted using several theoretical approaches. For example, Jayeola et al. (2020) have proposed the upper echelon theory as a framework to examine whether female- and male-led informal businesses differ in terms of performance. Lemma et al. (2023) have used liberal and social feminist theories to explain the performance gap between female-owned enterprises and their male-owned peers in Kenya and South Africa. Other perspectives include the social constructionist feminism (Justo et al., 2015), goal theory, theory of planned behavior, resource-based theory (Demartini, 2018; Watson et al., 2017), and entrepreneurship theory (Crane, 2022). The most commonly used theories are liberal and social feminist theories (Gottschalk & Niefert, 2013; Justo et al., 2015; Lemma et al., 2023; Westhead, 2003; Zolin et al., 2013). The liberal feminist theory suggests that women lack access to relevant resources, such as education, business experience, or financial capital. The social feminist theory suggests that women have different attitudes and values and adopt

different approaches to business (Arráiz, 2018; Bardasi et al., 2011; Boateng, 2018). Most studies have focused on female-owned firms instead of women-led businesses; that is, the primary interest is the performance of women-owned start-ups. The liberal and social feminist theories are compatible because liberal feminism focuses on gender barriers, whereas social feminism focuses on female behavior. Both theories may be unified into the social role theory (SRT), which considers two approaches to explaining gender roles in behavior and society.

A businessman's image is predominantly based on stereotypical notions of masculinity (Edelman et al., 2018). According to the SRT, gender differences and similarities in the behaviors of women and men are determined by gender roles (A. Eagly, 1987; A. H. Eagly & Mladinic, 1989; A. Eagly & Wood, 2016). Masculine gender roles include attributes, such as self-confidence, feeling superior, easily making decisions, being active, and being independent. Feminine gender roles include attributes like kindness, being helpful, being emotional, showing warmth to others, awareness of others' feelings, and gentleness (Abele, 2003; Spence & Helmreich, 1979). Previous research has demonstrated that feminine gender roles are strongly related to family roles, whereas masculine gender roles are associated with career success, business leadership, higher task orientation, and higher change potential (Abele, 2003; Kulich et al., 2018; Ramsey, 2017; Wille et al., 2018). These notions of expected female and male behaviors present the premise that women are less likely to succeed in the business world. For example, the traits desirable for a manager/entrepreneur are autonomy, need for achievement, self-efficacy, and risk-taking propensity (Comeig & Lurbe, 2018; Edelman et al., 2018; Rucker et al., 2018). These traits are closer to male stereotypes than to female gender roles. Therefore, it is plausible that women are less capable of achieving better firm performance because of the lack of these desirable attributes. Moreover, a businessowner's decision to hire a manager depends on the firm's goals and organizational design. Tondji (2022) has suggested that strategic delegation to overconfident managers, who have decided to overinvest in R&D and produce more market output, induces higher profits and welfare under certain conditions. Overconfidence is a male attribute; therefore, owners may prefer hiring male managers when R&D technology is less productive, and the main goal is cost reduction.

Our methodological review provides extensive evidence for the development of quantitative research on the FUH. As noted, the empirical evidence on this topic is mixed, and we found several studies without a theoretical foundation. In a study on top managers' gender differences in the context of firm performance in the manufacturing sector in the Czech Republic, Egerová

and Nosková (2019) have found that enterprises with women in top management teams have shown better financial performance than companies without women in top management teams. Conversely, Kristanti and Iswandi (2019), in a study of corporate firms in Indonesia, have suggested that although the influence of gender diversity is not significant, there are differences in a company's performance, when led by either female or male CEOs. Singhathep and Pholphirul (2015) have demonstrated that female general managers negatively affect short-term financial performance, including sales and annual benefits, in manufacturing firms in Thailand. Following the literature, we enunciate the FUH in Watson's (2002, p.92) words.

H1. Female-controlled businesses (on average) will generate lower outputs measured in terms of Return on Equity (ROE), Return on Assets (ROA), and net income per employee than male-controlled businesses.

Methodology

A two-stage approach is adopted in this study. First, a methodological review of previous empirical studies on the relationship between female leadership and firm performance is conducted. This review aims to identify the quantitative models and variables used to test FUH in female-managed firms. In the second stage, we specify three econometric models for testing the FUH based on Chilean business data, using the findings from the methodological review.

Methodological Review

We conducted a methodological review and selected the most representative empirical research on the relationship between female managers and firm performance. Our method is similar to previous methodological reviews in the management literature (Ball & Foster, 1982; Bash et al., 2021). Figure 1 shows the flowchart of the research design. We used the Web of Science (WoS) and Scopus databases to search for articles using the keywords "female underperformance hypothesis," "female CEO performance," "gender diversity board," and "female entrepreneur performance." We limit the results to publications from 2010 to the first half of 2020 because we detect increased research on this topic over these 10 years. Articles that did not contain keywords related to FUH during the first review were discarded. In the second review, we read the abstracts and selected studies with empirical analyses. We analyzed the introduction of each article in the third review stage. We thoroughly read each of the remaining articles to make the final selection for the fourth review.

We selected a final sample of 66 articles for quantitative analysis. Following an information-gathering formula, we grouped the quantitative elements of each article into the following categories: field of FUH, origin database, journal quartile classification, science category, sample selection criteria, methods, and main results. Among the 66 articles, we considered 26 that focused only on the relationship between female managers (e.g., CEO, CFO, TMT, or executive director) and firm performance. We obtained the variables and methods for testing the FUH from this process and used these outputs to select the methods and variables for empirical testing using our database.

Metrics in a FUH Empirical Research

A total of 66 quantitative articles in the WoS and Scopus databases published from 2010 to the first semester of 2020 were selected to test the FUH. We organized the papers' metrics by year, field (female managers, female owners, and board gender diversity), origin database, and journal quartile classification. The year with the highest number of publications was 2019 (11 from WoS and six from Scopus). Figure 2 shows the distribution of papers by year and the database of origin.

Tables 1 and 2 present information on the number of articles by the database of origin, journal quartile classification, scope of FUH, and sample type. Among the total articles, 24% corresponded to Scopus publications, and 76% were available in the WoS database. Most articles in Scopus were in the Q2 classification (41%), and more papers from WoS were in the Q1 classification (43%). Regarding article scope, 39% of the articles focused on the relationship between female managers and firm performance, 12% analyzed the effect of female ownership on firm performance, and 48% examined the influence of board gender diversity on firm financial outcomes. Regarding sample type, 68% of the articles used corporate firms, and 32% considered a mixed sample (corporate and non-corporate) from specific industries.

FUH research has been published in top journals, primarily in finance, management, business, economics, and entrepreneurship. Considering that FUH is a specialized field, we identified journals that have published two or more papers related to FUH across all themes (female manager, female owner, and board gender diversity). From Table 3, we can see that the *Journal of Business Ethics* has the majority of published articles (five papers).

Methods and Measures in a FUH Empirical Research

Our interest is to test FUH in female-managed businesses; hereafter, our analysis excludes FUH from the female businessowner and board gender diversity

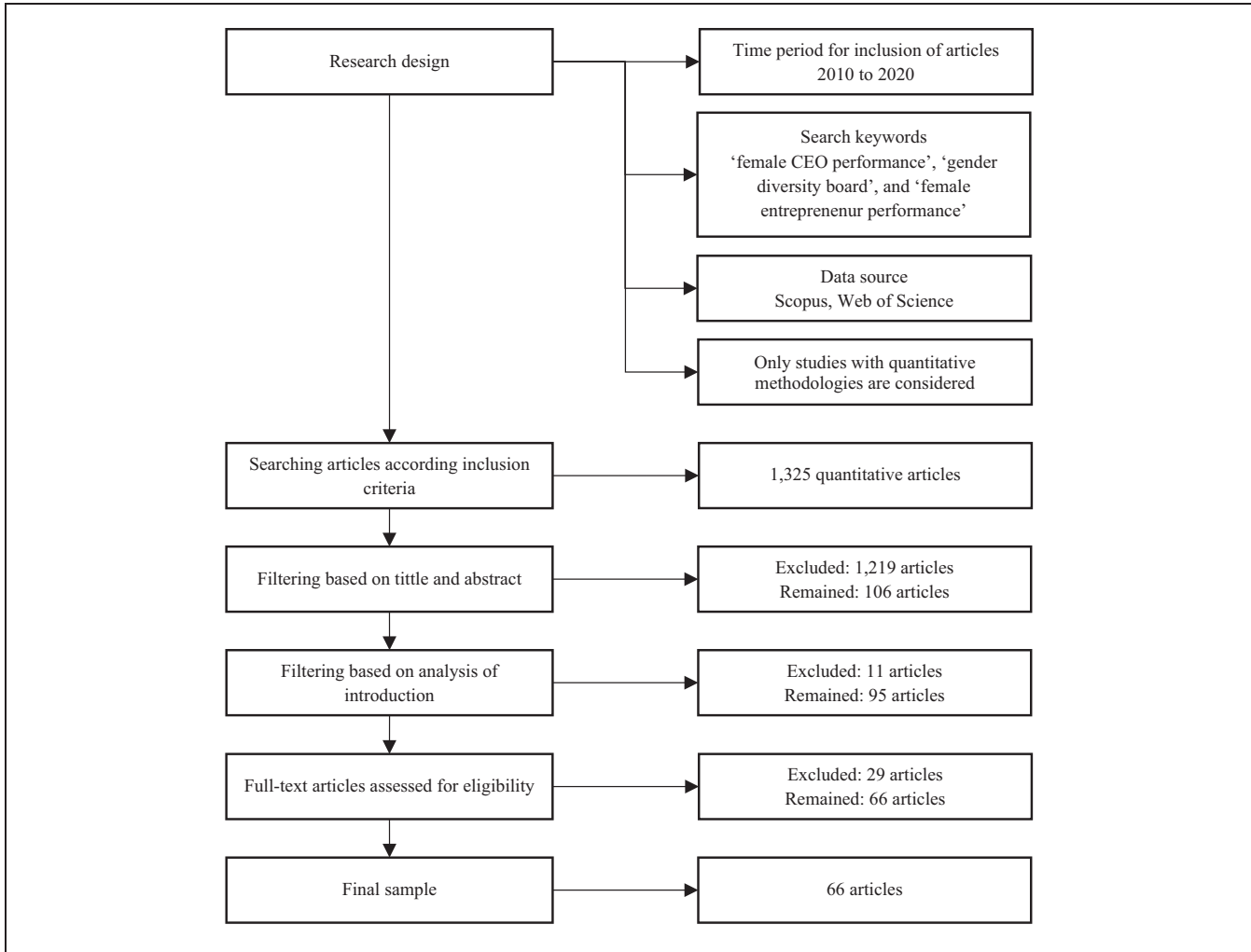


Figure 1. Literature review flowchart. Adapted from Hazaea et al. (2022).

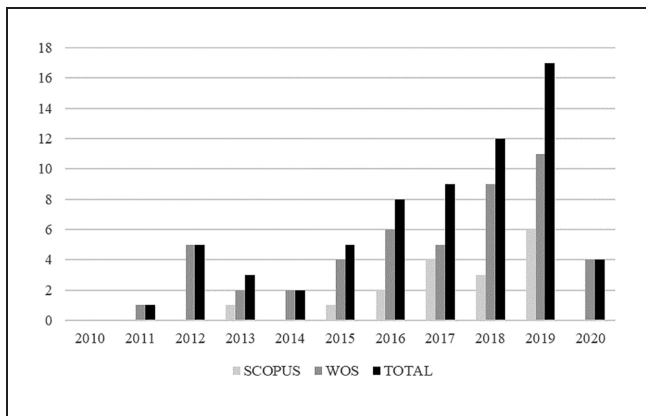


Figure 2. Distribution of papers by year and the database of origin. Metrics from June 2020. Source: Clarivate (2020) and SCImago & Scopus (2020).

context. Our focus on FUH in female managers is justified by a more comprehensive analysis of female leaders'

decision-making, which is unclear in the other two contexts (Strydom et al., 2017). Therefore, we identify 26 articles (out of 66 original) in which the authors tested the relationship between female managers and firm performance. Table 4 presents the 26 papers, indicating the author names, publication year, sample type, geographic context, methods, and main results.

The studies listed in Table 4 have different motivations for using specific methods. For example, Kolev (2012) has considered GLS to complement linear regression in studying CEOs' gender differences in corporate firm performance. This is because regression is blind to the fact that we might have one firm led by a female CEO in one period and 100 firms led by female CEOs in another period. Strøm et al. (2014) have built upon the Heckman (1978) endogenous dummy variable model and followed the two-step procedure in estimating managers' gender differences in the financial performance of microfinance institutions from different countries. Recently, to reduce selection bias because female

Table 1. Article Classification by Metrics and Source Per Year.

Year	Total Articles	Scopus Articles	WoS Articles	Scopus Quartile				WoS Quartile				
				Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	
2010	0	0	0									
2011	1	0	1									1
2012	5	0	5					2	2	1		
2013	3	1	2		1			1	1			
2014	2	0	2					2				
2015	5	1	4			1		3		1		
2016	8	2	6		1		1	1	3	2		
2017	9	4	5		2	1	1	2	1	1		1
2018	12	3	9		1	1	1	5	3			1
2019	17	6	11	2	2	1	1	3	2	3		3
2020	4	0	4					2	1			1
Total	66	17	49	2	7	4	4	21	13	8		7

Source. Clarivate (2020) and SCImago & Scopus (2020).

Note. Metrics at June 2020.

Table 2. Article Classification by Field and Sample Type Per Year.

Year	Field of FUH			Sample Type	
	Female Manager	Female Owner	Board Gender Diversity	Corporate	Non-Corporate
2010					
2011	1			1	
2012	3	2		2	3
2013	1	2		1	2
2014	2			1	1
2015	1	1	3	3	2
2016	3		5	7	1
2017	1	2	6	6	3
2018	4		8	11	1
2019	10		7	10	7
2020	0	1	3	3	1
Total	26	8	32	45	21

Source. Clarivate (2020) and SCImago & Scopus (2020).

Note. Metrics at June 2020.

Table 3. Journals With Two or More Articles in FUH.

Journal	Number of Articles	Database	JIF Quartile	Categories
<i>Journal of Business Ethics</i>	5	WoS	Q2	Ethics, Business
<i>Gender in Management</i>	4	WoS	Q4	Women's Studies, Management, Business
<i>Journal of Banking & Finance</i>	3	Wos	Q2	Economics, Business, Finance
<i>International Journal of Gender and Entrepreneurship</i>	3	Scopus	ESCI	Business and International Management, Economics and Finance, Gender Studies
<i>Journal of Business Research</i>	2	WoS	Q1	Business
<i>Journal of Business Venturing</i>	2	Wos	Q1	Business
<i>International Journal of Human Resource Management</i>	2	WoS	Q2	Management
<i>Leadership Quarterly</i>	2	WoS	Q1	Applied Psychology, Management
<i>Pacific-Basin Finance Journal</i>	2	WoS	Q2	Business, Finance

Source. Clarivate (2020) and SCImago & Scopus (2020).

Note. Metrics at June 2020. JIF = journal impact factor.

Table 4. Articles That Tested FUH in the Female Manager (CEO, CFO, or TMT) Context.

Author(s)/Year	Sample Type	Context	Methods	Main Results	Motivation for using the proposed method
Mkhethwa and Msweli (2011)	Corporate	South Africa	t-test	This study shows that listed South African companies with a high percentage of women in leadership positions do not outperform similar companies with a low percentage of women in leadership positions.	
Kolev (2012)	Corporate	S&P1000 Directory	OLS GLS	The research results show that female Chief Executive Officers underperform their male counterparts.	Regression is blind to the fact that in one period we might have one firm led by a female CEO, and in another period we might have 100 firms led by female CEOs. Therefore, a GLS model was used to complement the results.
Marco (2012)	Hospitality Industry	Spain	PFE	In ROA, women outperformed men, while in return on capital employed and profit per employee, there are no significant differences in profitability by gender.	“A linear regression model was estimated. The objective was consistent estimation of regression parameters given the existence of a latent explanatory variable, constant in time that may or may not be correlated with observable explanatory variables (p.986).”
Rodríguez-Domínguez et al. (2012)	Corporate	Spain	PFE PRE	The results obtained show that when working conditions and academic background are similar, women achieve better performance in sectors traditionally dominated by men.	“A dependence model based on a linear regression for panel data was selected as the analysis technique. More specifically, the models proposed were estimated through fixed and random effects, by checking the validity of the latter effects over the fixed ones using the Hausman test (p.611).”
Lam et al. (2013)	Corporate	China	OLS	Analysis of the CEO gender–firm performance association spawns mixed results.	
Liu et al. (2014)	Corporate	China	PFE IV	Results show that the female CEO has a positive effect on ROA.	“There are two estimation methods commonly used in the board and performance literature. One is the pooled ordinary least square (OLS) regression controlling for industry effects and the other is the panel regression with fixed effects. We apply the F-test to determine which is more appropriate for our study [...] we reject the pooled OLS approach in favor of the panel regression with fixed effects approach [based on F-test results] (p. 174).”

(continued)

Table 4. (continued)

Author(s)/Year	Sample Type	Context	Methods	Main Results	Motivation for using the proposed method
Strøm et al. (2014)	Microfinance Institutions	Multi-Country	2SLS	Results indicate that female-managed Microfinance Institutions have better performance than male-managed.	“We use a straightforward probit method to predict the female leadership variables... [This method] is fundamental in financial performance regressions where female leadership may be endogenously determined [...] In the financial performance estimations we build upon the Heckman (1978) endogenous dummy variable model and follow the two-step procedure (p. 65).”
Singhathep and Pholphirul (2015)	Manufacturing Firms	Thailand	OLS	Female general managers have a negative effect on short-term financial performance, including sales and annual benefits.	
Amore and Garofalo (2016)	Corporate Banks	United States	OLS	Results suggest that while banks with female executives experience significantly higher financial performance under low competition, they tend to underperform when competition increases.	
Reinert et al. (2016)	Banking Industry	Luxembourg	OLS	Results show a positive association between female management and firm performance.	“We provide evidence that FMS seems to have a direct impact on future firm performance, and that endogeneity issues (related to omitted variable bias and reverse causality) are not major concerns in our empirical analysis (p. 128).”
Perryman et al. (2016)	Corporate	Multi-Country	OLS	Results show that firms with greater gender diversity in TMTs have lower risk and deliver better performance.	
Ali and Shabir (2017)	Mixed	India	t-test	Analysis of firm performance across the gender of ownership finds that annual sales growth and labor productivity growth in female-owned enterprises is comparatively higher than that in male-owned enterprises.	“The difference in business performance and perceived business obstacles has been analyzed using the independent samples t-test. The independent samples t-test is the single most widely used test in statistics to compare differences between two group means (p. 7).”
Chadwick and Dawson (2018)	Corporate	S&P500	OLS	Results show a statistically significant and positive relationship between female leaders and firm performance only in nonfamily businesses.	

(continued)

Table 4. (continued)

Author(s)/Year	Sample Type	Context	Methods	Main Results	Motivation for using the proposed method
Nekhili et al. (2018)	Corporate	France	PSM	In family firms, the results show a positive and significant relationship between the appointment of a woman Chair and performance. In nonfamily firms, a negative and significant relationship is observed between the woman Chair and return on assets.	“A direct comparison of the performance between family firms and nonfamily firms is not very informative because performance might be explained by dissimilarities in characteristics between the two groups. To control for differences between family and nonfamily firms, we conduct a matched sample analysis using the propensity score matching (p.302).”
Martín-Ugedo et al. (2018)	Publishing Industry	Spain	2SLS 3SLS	Our results show that publishing companies whose CEO is female have higher performance than male-managed.	“The methodology employed is three stage least squares (3SLS). This methodology controls for the endogeneity of the variables, using a system of simultaneous equations. The alternative method to control for endogeneity, 2SLS presents consistent estimators but is not efficient (p. 116).”
Kanuri and Malm (2018)	Corporate	S&P1500	OLS	The results show that firms with female CEOs have higher average monthly and median returns, higher risk (standard deviation of monthly returns), and higher risk-adjusted performance (Sharpe, Sortino, and Omega ratios).	
Hoang et al. (2019)	Mixed	Vietnam	PFE	Female-managed firms have higher revenues and return on assets and capital than male-managed firms.	“To reduce the selection bias, we estimate a model using firm fixed-effects, which has the advantage that it eliminates the time-invariant unobserved variables (p. 128).”
Menicucci et al. (2019)	Hospitality Industry	Italia	PFE	When a regression model is designed to control other performance determinants (demographic, financial, and family variables), women-managed hotels outperform those managed by men for hotel growth.	“We estimated the regression parameters assuming the presence of a hidden explanatory variable (constant in time) that may or may not be associated with explanatory variables... We applied the fixed effects (FE) model as unit effects are not orthogonal to the explanatory variables (p. 630).”
Ullah et al. (2019)	Corporate	Pakistan	OLS	The findings illustrate that female CEO (FCEO) enhances a firm value.	“To find the impact of gender diversity (FDirectors and FCEOs) on firm value, we opt regression analysis following prior studies (p. 50)”

(continued)

Table 4. (continued)

Author(s)/Year	Sample Type	Context	Methods	Main Results	Motivation for using the proposed method
Kaur and Singh (2019)	Corporate	Nifty Top 500	PFE	The stated findings specify that long-tenured CEOs and firms led by female CEOs are negatively related to firm performance.	"In this specific regression specification, fixed effect model was accepted. This model stipulates that there are distinctive properties of individual firms, which are not formed by random variation and are also time invariant and, therefore, is an appropriate method in this research frame (p. 418)."
Flabbi et al. (2019)	Mixed	Italy	PFE	The impact of female leadership on firm performance increases with the share of female workers.	"The main challenge in estimating the impact of female CEOs on workers' wages and firms' performance is the sample selection bias induced by the non-random assignment of CEOs to firms. Our strategy to address these issues is to control for firm fixed effects, workforce composition effects, and CEO effects. (pp. 17–18)"
Ghosh and Guha (2019)	Microfinance Institutions	India	PFE PRE	The increase in female staff members leads to increased operational self-sufficiency and yield of the gross portfolio.	"The commonly used multivariate statistical data analysis technique for econometric, finance and social science research studies is panel data analysis as it can recognize the unobservable heterogeneity, which exists when the relationship between gender variables and performance variables are influenced by the unobserved factors... Random effects estimation is used to accommodate the influence of time-invariant explanatory variables (p.434)."
Jadiyappa et al. (2019)	Mixed	India	PFE PRE	The average ROA of the sample firms decreases by approximately 10% after a female enters the CEO role.	"We use the fixed effects estimator to estimate the coefficients. By using this estimator, we are able to control for the effects of the time-invariant, firm-specific factors which might affect the firm performance (p. 16)."

(continued)

Table 4. (continued)

Author(s)/Year	Sample Type	Context	Methods	Main Results	Motivation for using the proposed method
Beltran (2019)	Manufacturing and Service Firms	Multi-Country	OLS 2SLS	The findings suggest that a female owner strengthens the female CEO's business skills and leads to better firm performance than when the CEO is a woman and the owner is a man.	"It is important to consider the possibility that the explanatory variables of interest may not have the exogeneity condition necessary to ensure a consistent estimation of the parameters. Possible sources of endogeneity would be the bidirectional relation with the dependent variable, which could eventually introduce a bias in either direction (p.371)."
Kristanti and Iswandi (2019)	Corporate	Indonesia	PFE PRE	Although the influence of gender diversity was not significant, the different test results proved that there are differences in a company's performance, led by either female CEOs or male CEOs.	"This study used the data panel regression. To estimate the measurements, the Common Effect Model, the Fixed Effect Model and Random Effect Model were used. From all the models, the comparison was done using the Test of Chow and Hausman to get the most suitable model. The tests resulted in the Fixed Effect Model method as the most suitable method (p.244)."
Egerová and Nosková (2019)	Manufacturing Industry	Czech Republic	Cluster Analysis	The study found that, on average, enterprises with women in top management teams obtain better financial performance as measured by accounting-based measures such as ROA and ROE than companies without women in top management teams.	"To reveal basic relationships between indicators, the correlation matrix based on the calculation of Spearman's correlation coefficients was used. Namely, the k-means method was used. This method divides n observations into k clusters, where each observation belongs to the cluster with the nearest average (p. 133)."

Note. PSM = propensity score matching; 2SLS = two-stage least squares; 3SLS = three-stage least squares; OLS = ordinary least squares; PFE = panel fixed effects; PRE = panel random effects; IV = instrumental variables; CEO = chief executive officer; CFO = chief financial officer; TMT = top manager team.

leadership may be endogenously determined, Hoang et al. (2019) have specified a model using firm fixed effects to analyze managers' gender differences in Vietnamese firms across different economic sectors. However, 23% of these studies have not mentioned the reasons for their choice of method.

The main results of FUH studies reveal that only 23% of the findings support the claim that firms with female managers have lower performance than male-managed businesses. A positive relationship between female managers and firm performance has been observed in 42% of

studies. Finally, 35% of researchers have found mixed or non-significant results; that is, female-managed firms have better or lower performance than male-managed businesses depending on specific conditions, or insignificant relationships have been observed. The listed studies significantly differed in terms of the number of observations and database construction. Half of the studies used a sample of publicly listed corporations; these firms represent a small portion of enterprises in the business ecosystem. For example, Chile's Ministry of Economy has estimated that listed companies represent 2.5% of all

Chilean companies (INE, 2017). A similar phenomenon occurs in Spain, where less than 2% of Spanish companies are listed on stock exchanges (INE, 2022). Firms that do not trade publicly are very heterogeneous and subsequently better at embodying a variety of potential organizational forms and leadership styles (Mangematin et al., 2003). The lack of research in Latin American countries is noteworthy in the geographical context of testing FUH in female managers.

Ordinary least-squares regression (OLS) was the most frequently used methodology (35%). Only 23% of the studies used adequate methods to address endogeneity. A total of 10 studies used panel data and specified panel regression models with fixed effects or fixed plus random effects; only two of the studies used instrumental variables as endogenous regressors. The most common method for face endogeneity was instrumental variables (19%), specifically the two-stage least squares model (2SLS). One study specified a Propensity Score Matching model (PSM), taking gender as a selection variable (treatment) to estimate the differences in the performance of female- and male-managed firms. The most commonly used performance variables were ROE and ROA (65%); in four studies, the authors used an employee-based variable as a firm performance proxy (e.g., profit per employee, employee productivity, and value-added per employee). The main control variables were firm size and age, exports, family ownership, market share, and capital intensity. In all cases, the independent variable was the firm's manager's gender.

Testing Female Underperformance Hypothesis

Using the information collected from the methodological review, we selected three econometric methods, three dependent variables, and seven independent variables (including gender) from previous research on FUH in female-managed firms. We used data from the Fifth Longitudinal Survey of Firms (ELE5- *for its acronym in Spanish*). The last version, from 2017, contains information on 6,480 Chilean firms with different sizes and industries. We selected 2,323 companies with complete observations for each variable. Most firms do not trade publicly and are representative of the Chilean business ecosystem. We analyzed the data using the following three econometric models: OLS, 2SLS, and Matching Estimators. The criteria for including OLS and 2SLS are their high frequency of usage, and Matching Estimators are selected because this is a modern approach for solving endogeneity bias (see Diwisch et al., 2009; Simonsen & Skipper, 2006). The variables considered are those most commonly used in the literature on FUH in the context of female managers. The three dependent variables, ROE, ROA, and net income per employee (NIE), are proxies for the employee-based variable of firm performance. The

principal explanatory variable is the gender of the business manager, and the six control variables are average wages, capital intensity, whether the company exports, firm size and age, and market share.

The dependent variables for the three models were ROE, ROA, and NIE. We obtained ROE by dividing net income by total equity, ROA was calculated by dividing net income by total assets, and NIE by dividing net income by the total number of employees and applying a natural logarithm. The main explanatory variable (treatment) is *female manager*, a dummy variable that takes the value of one for companies with a female manager and zero otherwise. The control variables are *average wages*, obtained by dividing the firms' total paid salaries by the total number of employees in the natural logarithm; *capital intensity*, measured by dividing the total net assets by the total number of employees in the natural logarithm; *exports* is a dummy variable that takes the value one for firms that export and zero otherwise; *size* corresponds to the total number of workers in the company's natural logarithm; *age* is the number of years since the firm's founding, and *market share* is obtained by dividing each firm's net income by the total income of all firms in the sample. The data for each variable are annual cross-sections at the end of 2017.

OLS is the reference model, that is, the specification without considering endogeneity bias. We use 2SLS to obtain a consistent estimator of endogeneity (Maydeu-Olivares et al., 2019; Schmidt, 2020; Wooldridge, 2010). Endogenous explanatory variables can be instrumental based on the values these variables took in past periods (D. Chen et al., 2021; Hall, 1988; Stock & Watson, 2012; Y. Wang & Bellemare, 2019; Yogo, 2004). Therefore, we obtain instruments for the proposed endogenous variables (*average wage*, *capital intensity*, and *firm size*) by observing these variables for the same firms in the Fourth Longitudinal Survey of Firms (ELE4- *for its acronym in Spanish*) from 2015 (INE, 2015, 2017).

Non-parametric matching estimators are frequently used in impact evaluation studies (Abadie & Cattaneo, 2018; Clarke et al., 2019). The general idea behind this methodology is to determine the impact of treatment on outcomes using information on the treatment group and subjects similar to those in the treatment group who did not receive treatment. Using this information, we can construct a counterfactual for non-treatment (Lei & Candès, 2021; Vinha, 2006). We followed the following three approaches: nearest-neighbor matching with one neighbor, nearest-neighbor matching with five neighbors, and PSM. This approach allowed us to compare firms that are as similar as possible and whose main difference is the manager's gender, considering the same control variables proposed for the OLS and 2SLS models. We report the average treatment effect on the general population (ATE) and the average treatment effect on the

treated population (ATET). Overall, the ATE estimator is more rigorous than the ATET because the assumptions for ATET are less restrictive, and the standard error of the estimated ATET is generally larger than the standard error of the estimated ATE (Abadie et al., 2004; Imbens, 2004; Wooldridge, 2020).

Strategy to Approach Endogeneity

“Technically, endogeneity occurs when a predictor variable (x) in a regression model correlated with the error term (ϵ) in the model (Lynch & Brown, 2011, p. 112).” Concerns about endogeneity bias have recently increased because this specification problem is frequently encountered in econometric models (Cameron & Trivedi, 2005; Davidson & MacKinnon, 1993; Wooldridge, 2010, 2020). Moreover, models that use instrumental variables to deal with endogeneity bias have been refined to achieve better efficiency in econometric model estimation. The decision to consider an endogenous variable depends on the researcher and their knowledge of the theory underlying the research problem (Hamilton & Nickerson, 2003; Nakamura & Nakamura, 1998).

Suppose there is suspicion that any of the variables included in the model are endogenous. In this case, a variety of statistics allow us to test for an endogeneity specification problem. Our strategy for identifying the presence of endogenous variables includes a set of statistics that test for endogeneity, the weak-instrument problem, and overidentifying restrictions. If an endogeneity problem is not present in the model, the OLS estimator is more efficient than the model, including instrumental variables (Anatolyev & Skolkova, 2019). Therefore, it is essential to compare both estimators to evaluate which is more efficient (OLS vs. 2SLS). In this study, we use Hausman’s specification test to identify whether there is a systematic difference in the estimates; that is, the null hypothesis that the 2SLS estimators are indeed efficient (and consistent) estimators of the true parameters (Hausman, 1978).

Considering the evidence in this methodological review, we test endogeneity based on suggestions from previous studies and apply adequate statistics to assess this issue. We use the Durbin and Wu-Hausman statistics to test whether the suggested endogenous variables are exogenous. Both statistics test the null hypothesis that the proposed endogenous variables could be treated as exogenous. If both tests are significant, we can reject the null hypothesis of exogeneity and treat the variables under consideration as endogenous (Durbin, 1954; Hausman, 1978; Wu, 1974). Implementing a 2SLS estimator approach to deal with endogeneity bias is crucial for finding valid instruments; that is, variables that are sufficiently correlated with the included endogenous regressors but uncorrelated with the error term (Bound et al., 1995). Accordingly, we follow

the suggestions of Hall (1988), Stock and Watson (2012), and Yogo (2004) to use the endogenous variables lagged by one period as instruments.

We use the Anderson-Rubin test to test the hypothesis that the coefficients of the endogenous regressors in the structural equation are jointly equal to zero and that the overidentifying restrictions are valid (Anderson & Rubin, 1949; Baum et al., 2010). If the Anderson-Rubin test is significant, the null hypothesis of weak instruments can be rejected. We also report an F-version of the Cragg-Donald Wald statistic to test the hypothesis of weak instruments. If the Cragg-Donald test is significant, we can reject the null hypothesis of weak instruments (Cragg & Donald, 1993).

We report Sargan’s overidentification test of all instruments; however, our 2SLS equation is exactly identified; that is, we have the same number of instruments as the endogenous variables (Sargan, 1958, 1988). Similarly, we consider Anderson’s canonical correlation test to assess whether the instruments are correlated with the endogenous regressors; that is, if it satisfies the rank condition that the correlation or covariance between endogenous regressors and instruments is nonzero. If Anderson’s canonical correlation test is significant, we reject the null hypothesis of under-identification (Anderson, 1984; Baum et al., 2010).

The matching estimators “address the issue of self-selection bias and allow for a decomposition of treatment effects on outcomes (Titus, 2007, p. 487).” The FCEO variable may induce an endogeneity problem rooted in selection bias (self-selection) when estimating the differences in the performance of female- and male-led firms. Matching estimators (nearest-neighbor matching and PSM) deal with endogeneity bias because they allow treatment effects to be estimated by matching firms based on their similarities (Abadie et al., 2004; Abadie & Imbens, 2016). Therefore, the matching estimators isolate the effect of the third variables on the treatment effects. However, a simple matching estimator is biased when matching is not exact in finite samples (Abadie & Imbens, 2006, 2011). To reduce this bias, we include a bias adjustment based on the following covariates: *average wages*, *capital intensity*, *exports*, *size*, *age*, and *market share*. We follow the procedure proposed by Abadie et al. (2004) to estimate a bias-corrected matching estimator that adjusts the difference within the matches for differences in their covariate values. We use the Hausman specification test to determine which matching method generates consistent and efficient estimators using the nearest neighbor, with one neighbor as the reference model.

Results

Table 5 presents the overall sample descriptive statistics for each category of manager gender (men and women). Women run fewer than 20% of the firms in our sample.

Table 5. Descriptive Statistics.

Variable	Total sample		Firms run by women		Firms run by men	
	Mean	SD	Mean	SD	Mean	SD
ROE	0.397	0.898	0.363	0.953	0.406	0.884
ROA	0.038	0.702	-0.005	1.057	0.048	0.585
NIE	8.846	1.516	8.608	1.473	8.904	1.521
Female manager	0.194	0.396	-	-	-	-
Average wages	6.776	0.778	6.713	0.780	6.792	0.777
Capital intensity	9.051	1.932	8.804	1.977	9.111	1.917
Exports	0.177	0.381	0.164	0.371	0.179	0.384
Size	6.221	1.971	5.858	2.007	6.308	1.953
Age	21.690	14.258	20.465	13.707	21.985	14.376
Market share	4.31×10^{-4}	0.015	1.13×10^{-4}	0.001	0.001	0.017

Note. SD = standard deviation.

Table 6. Ordinary Least Square (OLS) Regression.

Variable	ROE		ROA		NIE	
	Coef.	SD	Coef.	SD	Coef.	SD
Female manager	-0.049	0.049	-0.052	0.046	-0.133***	0.047
Age	-0.003***	0.001	-0.001	0.001	-0.005***	0.001
Exports	-0.033	0.048	-0.012	0.009	0.145***	0.048
Average wages	-0.010	0.029	-0.015	0.050	0.506***	0.031
Capital intensity	-0.003	0.011	0.007	0.036	0.474***	0.014
Size	0.003	0.009	0.005	0.013	-0.045***	0.011
Market share	-0.044	0.115	-0.048	0.213	6.612***	0.552
Constant	0.565***	0.168	0.084	0.081	1.505***	0.191
R-Squared	0.004		0.002		0.627	
Observations	2,323		2,323		2,323	

Note. Estimate with robust standard errors. Coef. = Coefficient; SD = Standard deviation.

*** Significance level at 0.05/0.01.

The statistics by gender show that women-run firms have lower results for each of the result variables, and the remaining variables mirror this trend. Notable differences are observed in the firm variables of capital intensity and size. Women-led firms are smaller and less capital-intensive; that is, they use more work than capital in their operations (Recio, 1997). Low capital intensity could be motivated by the attitude of female management toward risk, with women tending to be more conservative in their decisions to invest capital (Loukil & Yousfi, 2016). Moreover, women-led businesses are concentrated in small and medium-size sectors (SMEs); therefore, women-led firms have fewer resources to invest in innovation, infrastructure, and business growth strategies (Guerrero et al., 2020; Ibáñez et al., 2020).

We use an OLS regression with robust errors for each of the performance variables (ROE, ROA, and NIE) as a reference model. The results are summarized in Table 6. Concerning the gender variable, in the OLS regressions, we observed that a female manager's presence negatively

and significantly affected the NIE variable. The relationship between gender and ROA/ROE was non-significant. We implemented an instrumental variable model, estimated using a two-stage least squares model, to deal with the supposed endogeneity problem. If the variables proposed as endogenous are exogenous, our reference model (OLS) produces estimators more efficiently than an alternative model that considers instrumental variables.

Tables 7 and 8 report the 2SLS results (first and second stages), including instrumental variables for dealing with endogeneity bias. Compared with the OLS model, the coefficients of the gender variable in 2SLS are non-significant for the three dependent variables (at the 0.05 significance level). Based on the results of the first-stage regressions, we reject the hypothesis that the matrix of reduced form coefficients has rank = K-1 (underidentified). Therefore, the Anderson canonical correlation test is highly significant, meaning that the instruments are sufficiently correlated with the endogenous variables and the rank condition is satisfied. The Cragg-Donald Wald

Table 7. First-Stage Instrumental Variables Regression.

Variables	Average wages		Capital intensity		Size		Market share	
	Coef.	SD	Coef.	SD	Coef.	SD	Coef.	SD
Female manager	-0.054*	0.032	-0.234***	0.088	-0.344***	0.091	1.6×10^{-4}	0.001
Age	0.002*	0.001	0.015***	0.002	0.025***	0.003	1.7×10^{-7}	1.5×10^{-5}
Exports	0.226***	0.034	0.275***	0.092	0.852***	0.096	-0.002***	0.001
IV Average wages	4.5×10^{-4} ***	1.3×10^{-5}	0.001***	3.6×10^{-5}	1.1×10^{-4} ***	3.7×10^{-5}	-4.6 × 10 ⁻⁷ **	2.2×10^{-7}
IV Capital intensity	1.3×10^{-9}	1.4×10^{-9}	2.6×10^{-8} ***	3.7×10^{-9}	-9.0 × 10 ⁻⁹ **	3.9×10^{-9}	-2.4 × 10 ⁻¹¹	2.3×10^{-11}
IV Size	9.4×10^{-7}	7.1×10^{-7}	-1.5 × 10 ⁻⁶	1.9×10^{-6}	3.4×10^{-5} ***	2.0×10^{-6}	-4.5 × 10 ⁻⁸ ***	1.2×10^{-8}
IV Market share	-4.364	3.503	44.494***	9.514	36.500***	9.895	3.069***	0.059
Constant	6.234***	0.028	7.817***	0.075	5.375***	0.078	1.1×10^{-4}	4.6×10^{-4}
Centered R ²	.372		.256		.219		.560	
Uncentered R ²	.992		.968		.931		.560	
Test of excluded instruments	297.81***		165.70***		93.10***		1.65	

Note. Prefix IV denotes instrumental variables. Coef.= Coefficient; SD= Standard deviation. * / ** / *** Significance level at 0.10/0.05/0.01.

statistic is significant (at the 0.05 significance level); therefore, we can reject the hypothesis of weak-instruments. These first-stage results are identical across the three proposed models (ROE, ROA, and NIE) because we use the same variables in all specifications. Considering that we included the same number of instruments for the endogenous regressor, the Sargan test of overidentification indicates that the equation is exactly identified for the three proposed models (ROE, ROA, and NIE).

The Anderson-Rubin test of joint significance of the endogenous regressors was significant only in the model with NIE as the dependent variable. Therefore, the endogenous regressors (jointly) do not significantly differ from zero in models with ROE- and ROA-dependent variables. Similarly, the Durbin score and Wu-Hausman test are significant only in the NIE model. Therefore, the proposed endogenous variables should be treated as exogenous in the ROE and ROA models. We compared the OLS and 2SLS estimators using the Hausman specification test. Consistent with our results given the absence of endogeneity in ROE and ROA models, we cannot reject the hypothesis that the OLS estimators are efficient (and consistent) estimators of the true parameters (ROE: $\chi^2 = 2.76$, $\rho > .05$; ROA: $\chi^2 = 0.42$, $\rho > .05$). Therefore, a comparison of the two estimators does not suggest substantial differences. In the NIE model, we reject the hypothesis of differences between OLS and 2SLS estimators; thus, the 2SLS estimators are efficient (and consistent) estimators of the true parameters (NIE: $\chi^2 = 12.69$, $\rho < .05$).

Briefly, a manager's gender does not influence firm performance because the relationship between female CEO and ROE/ROA is insignificant when considering the OLS and 2SLS estimators. Moreover, female-managed firms do not underperform compared to male-managed firms when estimations are treated for endogeneity bias in the NIE model.

The PSM and nearest-neighbor (with one and five neighbors) methods are used to estimate the difference in firm performance between female- and male-managed firms (Table 9). Non-significant differences are observed between the ROE of firms run by women- and male-managers, considering the ATE and ATET estimators of the three proposed models. The same results are observed for the ROA-dependent variable models. We find differences between the performance measures of NIE in the three models. When the PSM ATE estimator is used, female-managed firms show a significantly lower performance (NIE) than male-managed firms. However, when the nearest-neighbor with one-neighbor ATE and ATET estimators are considered, we do not find significant differences in NIE between female- and men-led businesses. Moreover, the PSM and nearest-neighbor with one-neighbor ATET estimators show non-significant

Table 8. Instrumental Variables Two-Stage Least Squares (2SLS) Regression.

Variable	ROE		ROA		NIE	
	Coef.	SD	Coef.	SD	Coef.	SD
Female manager	-0.067	0.051	-0.060	0.040	-0.094*	0.057
Age	-0.003	0.002	-0.001	0.002	-0.008***	0.002
Exports	-0.006	0.058	-0.005	0.045	0.153**	0.064
Average wages	0.017	0.179	0.003	0.141	0.063	0.198
Capital intensity	-0.009	0.089	-0.003	0.070	0.703***	0.099
Size	-0.033	0.031	-0.003	0.025	-0.021	0.035
Market share	1.844	2.478	0.217	1.943	5.694**	2.742
Constant	0.635	0.452	0.089	0.355	2.353***	0.500
Centered R-Squared	-.002		.001		.566	
Uncentered R-Squared	.169		.004		.988	
Anderson-Rubin test	0.38		0.01		143.54***	
Wu-Hausman test	0.569		0.035		4.105***	
Durbin (score)	2.288		0.139		16.389***	
p-value Sargan test	.000		.000		.000	
Anderson canonical test	39.08***		39.08***		39.08***	
Cragg-Donald test (weak ident.)	9.82**		9.82**		9.82**	
Observations	2,288		2,288		2,288	

Note. Prefix IV denotes instrumental variables. Coef.= Coefficient; SD= Standard deviation.

*/**/*** Significance level at 0.10/0.05/0.01.

Table 9. Estimation of Differences in Results by the Matching Methods.

Variables	Nearest-neighbor matching (1)		Nearest-neighbor matching (5)		Propensity-score matching	
	Coef.	SD	Coef.	SD	Coef.	SD
Average treatment effect (ATE)						
ROE	-0.072	0.055	-0.041	0.059	-0.017	0.052
ROA	-0.041	0.025	-0.030	0.019	-0.030	0.026
NIE	-0.007	0.134	0.025	0.171	-0.186***	0.134
Average treatment effect on the treated (ATET)						
ROE	-0.102	0.064	-0.090*	0.055	-0.073	0.070
ROA	-0.106	0.102	-0.060	0.056	-0.103	0.067
NIE	-0.106*	0.064	-0.150***	0.098	-0.164*	0.098
Treated obs	451		451		451	
Control obs	1,872		1,872		1,872	

Note. Nearest-neighbor matching (1): estimate with one neighbor. Nearest-neighbor matching (5): estimate with five neighbors. SD= Standard deviation.

*/**/*** Significance level at 0.10/0.05/0.01.

differences between female- and male-led firms. Female-managed firms show significantly lower performance (NIE) than male-managed firms when the nearest-neighbor with five-neighbor ATET estimator is used.

We implement the Hausman specification test using nearest-neighbor matching with one neighbor as the consistent estimator (reference model) to determine the best model for estimating the differences between female- and male-led business performances (Table 10). We test the null hypothesis that the difference in the coefficients is not systematic. We reject this hypothesis in all comparisons. The nearest-neighbors with five neighbors and PSM estimators are less efficient than nearest-neighbors with one neighbor estimator. Thus, the nearest neighbor with a one-

neighbor estimator is an efficient and consistent estimator of the true parameters. Therefore, female-managed firms do not underperform compared to male-managed firms.

Discussion

This research had the following two objectives: (1) to identify the empirical methods and most frequent variables in the literature to test the relationship between manager gender and firm performance and (2) to apply selected methods and variables for testing FUH using a Chilean business sample. We conducted a methodological review to assess and select empirical studies using

Table 10. Hausman Specification Test.

Models	ROE		ROA		NIE	
	χ^2	Prob > χ^2	χ^2	Prob > χ^2	χ^2	Prob > χ^2
Average treatment effect (ATE)						
Nearest-neighbor matching (5)	-0.30	-	0.14	.713	-0.02	-
Propensity-score matching	-45.37	-	0.13	.717	0.52	.470
Average treatment effect on the treated (ATET)						
Nearest-neighbor matching (5)	0.13	.717	0.24	.621	1.11	.292
Propensity-score matching	-1.02	-	0.00	.969	-0.45	-

Note. Nearest-neighbor matching (1): estimate with one neighbor. Nearest-neighbor matching (5): estimate with five neighbors. Nearest-neighbor matching (1) is the reference model to contrast (consistent estimator). SD= Standard deviation.

quantitative methods. We then conducted an empirical analysis to test the FUH for the following three models: OLS, nearest-neighbor matching, and PSM. We identified three dependent variables, the explanatory variable of manager gender, and the six control variables most commonly used in the literature. Additionally, our review revealed a significant gap in the research on female leadership in the Latin American context. We examined a sample that was more representative of the business ecosystem than the samples used in previous studies. This study shows no significant relationship between firms' financial performance and female business leadership when the endogeneity bias is resolved.

Conversely, other empirical research on FUH shows mixed results; several studies support the idea that a business run by women has lower financial performance than a firm managed by men (e.g., Lim et al., 2019; Singhathep & Pholphirul, 2015). Other studies have demonstrated a positive relationship between female managers and firm performance (e.g., Conyon & He, 2017; Moreno-Gómez & Calleja-Blanco, 2018). Additionally, some authors have argued that the manager's gender has no significant effect on firm performance or that this relationship depends on specific conditions (e.g., Singh et al., 2019; Unite et al., 2019).

There are several explanations for the variety of conclusions drawn from FUH empirical research, principally concerning methodological issues. Several studies have limited their samples to firms that trade in stock markets. However, this is a small portion of the business ecosystem and does not embody a variety of organizational forms and leadership structures (see: Valls & Cruz, 2019; Vu et al., 2019). Additionally, our methodological review showed that several studies did not correct for endogeneity or other specification problems. From the methodological review, we identified only 23% of the articles that recognized endogeneity bias and adopted adequate techniques to test for FUH in female management. Studies that have implemented models with endogeneity bias

solutions have shown no consensus on the effects of female managers on firm performance, that is, whether they have found positive, negative, or non-significant relations.

Economic models frequently have an endogeneity bias, in which the model's explanatory variables are strongly related to the outcome variables (Franzese, 2009; Nakamura & Nakamura, 1998). The instrumental variables approach (2SLS) is the most commonly used method for controlling endogeneity; however, obtaining valid instruments is difficult. An alternative solution is to instrumentalize the endogenous variables using lagged variables. Although this technique is not exempt from criticism, it is the most commonly used approach when better instruments are not available (see: Hall, 1988; Yogo, 2004). Another suitable method for addressing endogeneity is matching estimators. An important question to consider when comparing two groups (women vs. men) is how to differentiate between companies with specific characteristics. There is reason to believe that women-run firms differ from male-run firms in both observable and non-observable ways. In this case, non-parametric matching estimators (such as PSM) are helpful in comparing firms that are as similar as possible and whose main difference is rooted in the gender of their managers (e.g., Diwisch et al., 2009; Simonsen & Skipper, 2006). Previous research findings are inconclusive regarding the relationship between female management and firm performance (Brahma et al., 2021; Watson, 2020). According to Watson (2020), FUH is a myth. Our findings contribute to the theoretical and empirical confirmation of this claim, at least in relation to female management. We do not find relevant empirical evidence to support FUH when endogeneity bias is solved and a sample that is more representative of the business ecosystem is used.

We propose a research agenda that focuses on the positive role of women in business, instead of the assumption that female leaders negatively affect firm

performance. This new direction creates exciting opportunities for future research, as we leave behind the study of FUH. From a career development perspective, in which individuals try to climb the corporate ladder, it would be interesting to better understand how women perceive professional goals and the extent to which those goals represent their professional and personal expectations. Considering the differences in cultural contexts across countries, future research can explore how cultural stereotypes of masculinity might affect women's professional expectancies and ways to assess female managers' performance. This is an opportunity to conduct research in the Latin American context and emerging economies.

The extent to which female directors have decision-making autonomy is also a potential area of research. The mere presence of women on boards does not guarantee that they have the power to make decisions within the board; this is a weakness in policy design that imposes gender quotas on directive boards. Women's participation in board decision-making depends on the social interactions among board members and how women negotiate, acquire, and exercise power. Qualitative studies can provide new insights into the intricate dynamics of female leadership.

Conclusions

Our findings highlight the importance of using adequate econometric methods to establish the reliability and validity of the interpretations arising from quantitative analyses. Using an OLS model, we find that female managers have a significantly negative effect on NIE, but this result is not retained when 2SLS or matching estimators are implemented. We contribute to the research on female leadership in business by highlighting the methodological issues that must be considered in empirical studies. Additionally, our review reveals a significant gap in Latin American empirical research on female-led business performance and overall management, as pointed out earlier by many authors (Aguinis et al., 2020; Fritz & Silva, 2018; Nicholls-Nixon et al., 2011; Perez-Batres et al., 2012). Thus, we provide new evidence for female-managed firm performance in the Latin American context.

This study had a few limitations. These include the lack of longitudinal data, the absence of variables representing the professional capacities of managers, and the lack of structural aspects, such as business competency and dynamics. Business results have an intertemporal characteristic and probably do not match short-term decisions. Instead, they result from strategies that combine multi-dimensional objectives over different periods. Therefore, levels of education and managers' trajectories

can influence their ability to make decisions conducive to specific financial results (G. Wang et al., 2018). The composition of the industry, the intensity of competition, and the firms' position in the business context also determine firm performance. Finally, the benefits of gender diversity and female participation in leadership positions should not be evaluated solely based on an economic perspective (Hossain et al., 2017; McGuinness et al., 2017). Instead, we must ask how diversifying human capital, such as gender, racial-ethnic identification, and disability status, can affect business outcomes beyond financial indicators and profits.

Implications

This study had several theoretical, methodological, and practical implications. We show the roots and evidence of FUH testing based on previous research analyses, considering the most frequently used approaches. Theoretically, our results imply that the relationship between manager gender and firm performance is a context-based phenomenon because it depends on several factors, which may explain why we observe different results in different countries. From a methodological perspective, we expose the relevance of a well-specified model to test the relationship between manager gender and firm performance because we demonstrate the differences in the consistency and efficiency of estimators under various methods. Therefore, we recommend verifying whether an endogeneity bias exists and the quality and adequacy of the sample before running the model.

FUH is also present in shared knowledge; in practice, people tend to believe that women are less capable of successfully running a business. This notion is rooted in social gender hierarchies that may differ across countries. Therefore, public policymakers should work to reduce gender gaps in business and other social spheres, emphasizing the legitimization of women in top leadership positions. Considering that the negative relationship between female managers and firm performance is not always true, practitioners may consider hiring more women in top management positions, given the other benefits that women's leadership may have in overall firm performance and not only in the financial results.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article:

This research was funded by ANID PIA/Basal FB0002; the ANID/FONDAP/15130015.

ORCID iDs

María José Ibáñez  <https://orcid.org/0000-0002-5221-1569>
Roberto D. Ponce Oliva  <https://orcid.org/0000-0001-5024-7456>

Data Availability Statement

Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

References

- Abadie, A., & Cattaneo, M. D. (2018). Econometric methods for program evaluation. *Annual Review of Economics*, 10(1), 465–503. <https://doi.org/10.1146/annurev-economics-080217-053402>
- Abadie, A., Drukker, D., Herr, J. L., & Imbens, G. W. (2004). Implementing matching estimators for average treatment effects in Stata. *The Stata Journal: Promoting Communications on Statistics and Stata*, 4(3), 290–311. <https://doi.org/10.1177/1536867x0400400307>
- Abadie, A., & Imbens, G. W. (2006). Large sample properties of matching estimators for average treatment effects. *Econometrica*, 74(1), 235–267. <https://doi.org/10.1111/j.1468-0262.2006.00655.x>
- Abadie, A., & Imbens, G. W. (2011). Bias-corrected matching estimators for average treatment effects. *Journal of Business and Economic Statistics*, 29(1), 1–11. <https://doi.org/10.1198/jbes.2009.07333>
- Abadie, A., & Imbens, G. W. (2016). Matching on the estimated propensity score. *Econometrica*, 84(2), 781–807. <https://doi.org/10.3982/ecta11293>
- Abele, A. E. (2003). The dynamics of masculine-agentic and feminine-communal traits: Findings from a prospective study. *Journal of Personality and Social Psychology*, 85(4), 768–776. <https://doi.org/10.1037/0022-3514.85.4.768>
- Aggarwal, R., Jindal, V., & Seth, R. (2019). Board diversity and firm performance: The role of business group affiliation. *International Business Review*, 28(6), 101600. <https://doi.org/10.1016/j.ibusrev.2019.101600>
- Aguinis, H., Villamor, I., Lazzarini, S. G., Vassolo, R. S., Amorós, J. E., & Allen, D. G. (2020). Conducting Management Research in Latin America: Why and what's in it for You? *Journal of Management*, 46(5), 615–636. <https://doi.org/10.1177/0149206320901581>
- Ahl, H. (2006). Why research on women entrepreneurs needs new directions. *Entrepreneurship Theory and Practice*, 30(5), 595–621.
- Ahl, H., & Marlow, S. (2012). Exploring the dynamics of gender, feminism and entrepreneurship: advancing debate to escape a dead end? *Organization*, 19(5), 543–562. <https://doi.org/10.1177/1350508412448695>
- Ali, J., & Shabir, S. (2017). Does gender make a difference in business performance? *Gender in Management An International Journal*, 32(3), 218–233. <https://doi.org/10.1108/gm-09-2016-0159>
- Amore, M. D., & Garofalo, O. (2016). Executive gender, competitive pressures, and corporate performance. *Journal of Economic Behavior & Organization*, 131, 308–327. <https://doi.org/10.1016/j.jebo.2016.09.009>
- Anatolyev, S., & Skolkova, A. (2019). Many instruments: Implementation in Stata. *The Stata Journal: Promoting Communications on Statistics and Stata*, 19(4), 849–866. <https://doi.org/10.1177/1536867x19893627>
- Anderson, T. W. (1984). *Introduction to multivariate statistical analysis (second)*. John Wiley & Sons.
- Anderson, T. W., & Rubin, H. (1949). Estimation of the parameters of a single equation in a complete system of stochastic equations. *Annals of Mathematical Statistics*, 20(1), 46–63. <https://doi.org/10.1214/aoms/1177730090>
- Arráiz, I. (2018). Time to share the load: Gender differences in household responsibilities and business profitability. *Small Business Economics*, 51(1), 57–84. <https://doi.org/10.1007/s11187-017-9925-z>
- Ball, R., & Foster, G. (1982). Corporate financial reporting: A methodological review of empirical research. *Journal of Accounting Research*, 20, 161. <https://doi.org/10.2307/2674681>
- Bardasi, E., Sabarwal, S., & Terrell, K. (2011). How do female entrepreneurs perform? Evidence from three developing regions. *Small Business Economics*, 37(4), 417–441. <https://doi.org/10.1007/s11187-011-9374-z>
- Bash, K. L., Howell Smith, M. C., & Trantham, P. S. (2021). A systematic methodological review of hierarchical linear modeling in mixed methods research. *Journal of Mixed Methods Research*, 15(2), 190–211. <https://doi.org/10.1177/1558689820937882>
- Baum, C. F., Schaffer, M. E., & Stillman, S. (2010). *ivreg2: Stata module for extended instrumental variables/2SLS, GMM and AC HAC, LIML and k-class regression*. <http://ideas.repec.org/c/boc/bocode/s425401.html>
- Bear, S., Rahman, N., & Post, C. (2010). The impact of board diversity and gender composition on corporate social responsibility and firm reputation. *Journal of Business Ethics*, 97(2), 207–221. <https://doi.org/10.1007/s10551-010-0505-2>
- Beltran, A. (2019). Female leadership and firm performance. *Prague Economic Papers*, 28(3), 363–377. <https://doi.org/10.18267/j.pep.695>
- Bennouri, M., Chtioui, T., Nagati, H., & Nekhili, M. (2018). Female board directorship and firm performance: What really matters? *Journal of Banking & Finance*, 88, 267–291. <https://doi.org/10.1016/j.jbankfin.2017.12.010>
- Boateng, A. (2018). Contextualising women's entrepreneurship in Africa. In A. Boateng (Ed.), *African female entrepreneurship: Merging profit and social motives for the greater good* (pp. 3–33). Springer International Publishing.
- Bound, J., Jaeger, D. A., & Baker, R. M. (1995). Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *Journal of the American Statistical Association*, 90(430), 443. <https://doi.org/10.2307/2291055>
- Brahma, S., Nwafor, C., & Boateng, A. (2021). Board gender diversity and firm performance: The UK evidence. *International Journal of Finance & Economics*, 26(4), 5704–5719. <https://doi.org/10.1002/ijfe.2089>

- Brush, C. (2019). Growth-oriented women entrepreneurs: Strategies for raising money. In V. L. Crittenden (Ed.), *Go-to-market strategies for women entrepreneurs* (pp. 271–281). Emerald Publishing Limited.
- Brush, C. G. (1992). Research on women business owners: Past trends, a new perspective and future directions. *Entrepreneurship Theory and Practice*, 16(4), 5–30. <https://doi.org/10.1177/104225879201600401>
- Cameron, A. C., & Trivedi, P. K. (2005). *Microeconometrics: Methods and applications*. Cambridge University Press.
- Carmona, S., Ezzamel, M., & Mogotocoro, C. (2018). Gender, management styles, and forms of capital. *Journal of Business Ethics*, 153(2), 357–373. <https://doi.org/10.1007/s10551-016-3371-8>
- Chadwick, I. C., & Dawson, A. (2018). Women leaders and firm performance in family businesses: An examination of financial and nonfinancial outcomes. *Journal of Family Business Strategy*, 9(4), 238–249. <https://doi.org/10.1016/j.jfbs.2018.10.002>
- Chaganti, R. (1986). Management in women-owned enterprises. *Journal of Small Business Management*, 24(4), 18–29.
- Chen, D., Chen, J., & Li, S. (2021). Instrumental variable quantile regression of spatial dynamic Durbin panel data model with fixed effects. *Mathematics*, 9(24), 3261. <https://doi.org/10.3390/math9243261>
- Chen, J., Leung, W. S., & Evans, K. P. (2018). Female board representation, corporate innovation and firm performance. *Journal of Empirical Finance*, 48, 236–254. <https://doi.org/10.1016/j.jempfin.2018.07.003>
- Christiansen, L. E., Lin, H., Pereira, J., Topalova, P., & Turk, R. (2016). Gender diversity in senior positions and firm performance: Evidence from Europe. *IMF Working Papers*, 16(50), 1–29. <https://doi.org/10.5089/9781513553283.001>
- Clarivate. (2020). *Journal Citation Reports*. InCites Journal Citation Reports. <https://jcr-clarivate-com.suscripciones.udd.cl:2443/JCRLandingPageAction.action>
- Clarke, G. M., Conti, S., Wolters, A. T., & Steventon, A. (2019). Evaluating the impact of healthcare interventions using routine data. *BMJ*, 365(12239), 12239–NaN7. <https://doi.org/10.1136/bmj.12239>
- Comeig, I., & Lurbe, M. (2018). Gender behavioral issues and entrepreneurship. In A. Tur-Porcar & D. Ribeiro-Soriano (Eds.), *Inside the mind of the entrepreneur. Contributions to management science* (pp. 149–159). Springer.
- Connell, C. M. (2019). Women CEOs on making strategy happen. *Strategic Direction*, 35(7), 1–4. <https://doi.org/10.1108/sd-09-2018-0184>
- Canyon, M. J., & He, L. (2017). Firm performance and boardroom gender diversity: A quantile regression approach. *Journal of Business Research*, 79, 198–211. <https://doi.org/10.1016/j.jbusres.2017.02.006>
- Cook, S. L., & Goodman, L. A. (2006). Beyond frequency and severity: Development and validation of the brief coercion and conflict scales. *Violence Against Women*, 12(11), 1050–1072. <https://doi.org/10.1177/1077801206293333>
- Cragg, J. G., & Donald, S. G. (1993). Testing identifiability and specification in instrumental variable models. *Economic Theory*, 9(2), 222–240. <https://doi.org/10.1017/s0266466600007519>
- Crane, S. R. (2022). Entrepreneurship and economic growth: Does gender matter? *International Journal of Gender and Entrepreneurship*, 14(1), 3–25. <https://doi.org/10.1108/ijge-04-2021-0056>
- Crittenden, V., & Bliton, K. (2019). Direct selling: The power of women helping women. In V. L. Crittenden (Ed.), *Go-to-market strategies for women entrepreneurs* (pp. 195–205). Emerald Publishing Limited.
- Cuba, R., Decenzo, D., & Anish, A. (1983). Management practices of successful female business owners. *American Journal of Small Business*, 8(2), 40–46. <https://doi.org/10.1177/104225878300800208>
- Đặng, R., Houanti, L., Reddy, K., & Simioni, M. (2020). Does board gender diversity influence firm profitability? A control function approach. *Economic Modelling*, 90, 168–181. <https://doi.org/10.1016/j.econmod.2020.05.009>
- Davidson, R., & MacKinnon, J. G. (1993). *Estimation and inference in econometrics*. Oxford University Press.
- Dean, H., Larsen, G., Ford, J., & Akram, M. (2019). Female Entrepreneurship and the metanarrative of economic growth: A critical review of underlying assumptions. *International Journal of Management Reviews*, 21(1), 24–49. <https://doi.org/10.1111/ijmr.12173>
- Demartini, P. (2018). Innovative female-led startups. Do women in business underperform? *Administrative Sciences*, 8(4), 70. <https://doi.org/10.3390/admsci8040070>
- Diwisch, D. S., Voithofer, P., & Weiss, C. R. (2009). Succession and firm growth: Results from a non-parametric matching approach. *Small Business Economics*, 32(1), 45–56. <https://doi.org/10.1007/s11187-007-9072-z>
- Durbin, J. (1954). Errors in variables. *Revue de l'Institut International de Statistique | Review of the International Statistical Institute*, 22(1/3), 23. <https://doi.org/10.2307/1401917>
- Du Rietz, A., & Henrekson, M. (2000). Testing the female underperformance hypothesis. *Small Business Economics*, 14(1), 1–10. <https://doi.org/10.1023/a:1008106215480>
- Eagly, A. (1987). *Sex differences in social behavior: A social-role interpretation*. Erlbaum.
- Eagly, A. H., & Mladinic, A. (1989). Gender stereotypes and attitudes toward women and men. *Personality and Social Psychology Bulletin*, 15(4), 543–558. <https://doi.org/10.1177/0146167289154008>
- Eagly, A., & Wood, W. (2016). Social role theory of sex differences. In A. Wong, M. Wickramasinghe, R. Hoogland, & N. A. Naples (Eds.), *The Wiley Blackwell Encyclopedia of Gender and Sexuality Studies*. John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781118663219.wbegss183>
- Edelman, L. F., Donnelly, R., Manolova, T., & Brush, C. G. (2018). Gender stereotypes in the angel investment process. *International Journal of Gender and Entrepreneurship*, 10(2), 134–157. <https://doi.org/10.1108/ijge-12-2017-0078>
- Egerová, D., & Nosková, M. (2019). Top management team composition and financial performance: Examining the role of gender diversity. *E + M Ekonomie a Management*, 22(2), 129–143. <https://doi.org/10.15240/tul/001/2019-2-009>
- Ellemers, N., & Nadal, K. (2018). Gender stereotypes. *Annual Review of Psychology*, 69(1), 275–298. <https://doi.org/10.1146/annurev-psych-122216-011719>

- Fairlie, R. W., & Robb, A. M. (2009). Gender differences in business performance: Evidence from the characteristics of Business Owners survey. *Small Business Economics*, 33(4), 375–395. <https://doi.org/10.1007/s11187-009-9207-5>
- Farhan, F., & Nayan, S. (2018). An empirical evidence on the effect of women board representation on firm performance of companies listed in Iraq Stock Exchange. *Business and Economic Horizons*, 14(1), 117–131.
- Fischer, E. (1992). Sex differences and small-business performance among Canadian retailers and service providers. *Journal of Small Business & Entrepreneurship*, 9(4), 2–13. <https://doi.org/10.1080/08276331.1992.10600408>
- Flabbi, L., Macis, M., Moro, A., & Schivardi, F. (2019). Do female executives make a difference? The impact of female leadership on gender gaps and firm performance. *Econometrics Journal*, 129(622), 2390–2423. <https://doi.org/10.1093/ej/uez012>
- Foley, S., Hang-Yue, N., & Wong, A. (2005). Perceptions of discrimination and justice: Are there gender differences in outcomes? *Group & Organization Management*, 30(4), 421–450. <https://doi.org/10.1177/1059601104265054>
- Franzese, R. J. (2009, 2 September). Multicausality, context-conditional, and endogeneity. In C. Boix & S. C. Stokes (Eds.), *The Oxford handbook of comparative politics*. Oxford Academic. <https://doi.org/10.1093/oxfordhb/9780199566020.003.0002>
- Fritz, M. M. C., & Silva, M. E. (2018). Exploring supply chain sustainability research in Latin America. *International Journal of Physical Distribution & Logistics Management*, 48(8), 818–841. <https://doi.org/10.1108/ijpdlm-01-2017-0023>
- Ghosh, C., & Guha, S. (2019). Role of gender on the performance of Indian microfinance institutions. *Gender in Management An International Journal*, 34(6), 429–443. <https://doi.org/10.1108/gm-03-2019-0036>
- Gipson, A. N., Pfaff, D. L., Mendelsohn, D. B., Catenacci, L. T., & Burke, W. W. (2017). Women and Leadership: Selection, development, leadership style, and performance. *Journal of Applied Behavioral Science*, 53(1), 32–65. <https://doi.org/10.1177/0021886316687247>
- Gottschalk, S., & Niefert, M. (2013). Gender differences in business success of German start-up firms. *International Journal of Entrepreneurship and Small Business*, 18(1), 15–46.
- Green, C. P., & Homroy, S. (2018). Female directors, board committees and firm performance. *European Economic Review*, 102, 19–38. <https://doi.org/10.1016/j.euroecorev.2017.12.003>
- Grover, V. (2015). Second generation gender bias : Invisible barriers holding women back in organizations. *International Journal of Applied Research*, 1(4), 1–4.
- Guerrero, M., Serey, T., Ibáñez, M. J., Romani, G., & Fernandez, C. (2020). Mujeres y actividad emprendedora Chile 2019. *GEM Report*, 1(1), 1–83.
- Gupta, V. K., Wieland, A. M., & Turban, D. B. (2019). Gender characterizations in Entrepreneurship: A multi-level investigation of sex-role stereotypes about High-Growth, commercial, and Social Entrepreneurs. *Journal of Small Business Management*, 57(1), 131–153. <https://doi.org/10.1111/jsbm.12495>
- Hall, R. E. (1988). Intertemporal substitution in consumption. *Journal of Political Economy*, 96(2), 339–357. <https://doi.org/10.1086/261539>
- Hamilton, B. H., & Nickerson, J. A. (2003). Correcting for endogeneity in strategic management research. *Strategic Organization*, 1(1), 51–78. <https://doi.org/10.1177/1476127003001001218>
- Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica*, 46(6), 1251. <https://doi.org/10.2307/1913827>
- Hazaea, S. A., Zhu, J., Khatib, S. F. A., Bazhair, A. H., & Elamer, A. A. (2022). Sustainability assurance practices: A systematic review and future research agenda. *Environmental Science and Pollution Research*, 29(4), 4843–4864. <https://doi.org/10.1007/s11356-021-17359-9>
- Heckman, J. J. (1978). Simple statistical models for discrete panel data developed and applied to test the hypothesis of true state dependence against the hypothesis of spurious state dependence. *Annales de l'insée*, 30 (31), 227–269. <https://doi.org/10.2307/20075292>
- Heilman, M. E. (2012). Gender stereotypes and workplace bias. *Research in Organizational Behavior*, 32, 113–135. <https://doi.org/10.1016/j.riob.2012.11.003>
- Helfat, C., Harris, D., & Wolfson, P. (2006). The pipeline to the top: Women and men in the top executive ranks of U.S. Corporations. *Academy of Management Perspectives*, 20(4), 42–64. <https://doi.org/10.5465/AMP.2006.23270306>
- Hisrich, R., & Brush, C. (1987). *Women business owners: A longitudinal study* (pp. 21–39). Frontiers of Business Ownership Research, Center for Entrepreneurial Studies, Babson College.
- Hisrich, R., & Brush, C. (2019). The woman entrepreneur: Management skills and business problems. *Journal of Small Business Management*, 57(1), 30–37.
- Hoang, T., Nguyen, C., & Phung, T. (2019). Do male CEOs really run firms better than female counterparts? New evidence from vietnam. *Hitotsubashi Journal of Economics*, 60(2), 121–140.
- Hossain, M., Farooque, O. A., Momin, M. A., & Almotairy, O. (2017). Women in the boardroom and their impact on climate change related disclosure. *Social Responsibility Journal*, 13(4), 828–855. <https://doi.org/10.1108/srj-11-2016-0208>
- Ho, S. S. M., Li, A. Y., Tam, K., & Zhang, F. (2015). CEO gender, ethical leadership, and accounting conservatism. *Journal of Business Ethics*, 127, 351–370. <https://doi.org/10.1007/s10551-013-2044-0>
- Hundley, G. (2001). Why women earn less than men in self-employment. *Journal of Labor Research*, 22(4), 817–829. <https://doi.org/10.1007/s12122-001-1054-3>
- Ibáñez, M. J., Guerrero, M., & Mahto, R. V. (2020). Women-led SMEs: Innovation and collaboration → performance? *Journal of the International Council for Small Business*, 1(3-4), 111–117. <https://doi.org/10.1080/26437015.2020.1850155>
- ILO Bureau for Employers' Activities. (2020). *Improving gender diversity in company boards*.
- Imbens, G. W. (2004). Nonparametric estimation of average treatment effects under exogeneity: A review. *Review of Economics and Statistics*, 86(1), 4–29. <https://doi.org/10.1162/003465304323023651>

- INE. (2015). *Cuarta Encuesta Longitudinal de Empresas*. Ministerio de Economía. <https://www.economia.gob.cl/2017/03/16/cuarta-encuesta-longitudinal-de-empresas-ele4.htm>
- INE. (2017). *Quinta Encuesta Longitudinal de Empresas*. Ministerio de Economía.
- INE. (2022). *Estadística de sociedades mercantiles. Últimos datos*. INEbase. https://ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736177026&menu=ulti-Datos&idp=1254735576550
- Isidro, H., & Sobral, M. (2015). The effects of women on corporate boards on firm value, financial performance, and ethical and social compliance. *Journal of Business Ethics*, 132(1), 1–19. <https://doi.org/10.1007/s10551-014-2302-9>
- Jadiyappa, N., Jyothi, P., Sireesha, B., & Hickman, L. E. (2019). CEO gender, firm performance and agency costs: Evidence from India. *Journal of Economic Studies*, 46(2), 482–495. <https://doi.org/10.1108/jes-08-2017-0238>
- Jayeola, O., Sidek, S., Owoye, I., & Kazeem, Y. K. (2020). Gender and the performance of informal sector enterprises. *European Scientific Journal*, 16(4), 57. <https://doi.org/10.19044/esj.2020.v16n4p57>
- Johnson, S., & Storey, D. (1993). Male and female entrepreneurs and their businesses. In S. Allen & C. Truman (Eds.), *Women in business: Perspectives on women entrepreneurs* (pp. 70–85). Routledge.
- Justo, R., DeTienne, D. R., & Sieger, P. (2015). Failure or voluntary exit? Reassessing the female underperformance hypothesis. *Journal of Business Venturing*, 30(6), 775–792. <https://doi.org/10.1016/j.jbusvent.2015.04.004>
- Kanuri, S., & Malm, J. (2018). Performance of female CEOs. *Journal of Investing*, 27(1), 135–142. <https://doi.org/10.3905/joi.2018.27.1.135>
- Kaur, R., & Singh, B. (2019). Do CEO characteristics explain firm performance in India? *Journal of Strategy and Management*, 12(3), 409–426. <https://doi.org/10.1108/jsma-02-2019-0027>
- Kirsch, A. (2018). The gender composition of corporate boards: A review and research agenda. *Leadership Quarterly*, 29(2), 346–364. <https://doi.org/10.1016/j.leaqua.2017.06.001>
- Kolev, G. I. (2012). Underperformance by female CEOs: A more powerful test. *Economics Letters*, 117(2), 436–440. <https://doi.org/10.1016/j.econlet.2012.06.028>
- Kotiranta, A., Kovalainen, A., & Rouvinen, P. (2010). Chapter 4: Female leadership and company profitability. In C. G. Brush, A. Bruin, E. J. Gatewood, & C. Henry (Eds.), *Women entrepreneurs and the global environment for growth*. Edward Elgar Publishing. <https://doi.org/10.4337/9781849806633.00009>
- Krishnan, H. A., & Park, D. (2005). A few good women—on top management teams. *Journal of Business Research*, 58(12), 1712–1720. <https://doi.org/10.1016/j.jbusres.2004.09.003>
- Kristanti, F. T., & Iswandi, I. (2019). The differences of company's performance from CEO diversity. *Polish Journal of Management Studies*, 19(2), 240–249. <https://doi.org/10.17512/pjms.2019.19.2.20>
- Kulich, C., Iacoviello, V., & Lorenzi-Cioldi, F. (2018). Solving the crisis: When agency is the preferred leadership for implementing change. *Leadership Quarterly*, 29(2), 295–308. <https://doi.org/10.1016/j.leaqua.2017.05.003>
- Lam, K. C. K., McGuinness, P. B., & Vieito, J. P. (2013). CEO gender, executive compensation and firm performance in Chinese-listed enterprises. *Pacific-Basin Finance Journal*, 21(1), 1136–1159. <https://doi.org/10.1016/j.pacfin.2012.08.006>
- Lei, L., & Candès, E. J. (2021). Conformal inference of counterfactuals and individual treatment effects. *Journal of the Royal Statistical Society Series B (Statistical Methodology)*, 83(5), 911–938. <https://doi.org/10.1111/rssb.12445>
- Lemma, T. T., Gwatidzo, T., & Mlilo, M. (2023). Gender differences in business performance: Evidence from Kenya and South Africa. *Small Business Economics*, 60, 591–614. <https://doi.org/10.1007/s11187-022-00605-w>
- Lim, K. P., Lye, C.-T., Yuen, Y. Y., & Teoh, W. M. Y. (2019). Women directors and performance: Evidence from Malaysia. *Equality Diversity and Inclusion An International Journal*, 38(8), 841–856. <https://doi.org/10.1108/edi-02-2019-0084>
- Liu, Y., Wei, Z., & Xie, F. (2014). Do women directors improve firm performance in China? *Journal of Corporate Finance*, 28, 169–184. <https://doi.org/10.1016/j.jcorpfin.2013.11.016>
- Loscocco, K. A., Robinson, J., Hall, R. H., & Allen, J. K. (1991). Gender and small business success: An inquiry into women's relative disadvantage. *Social Forces*, 70(1), 65–85. <https://doi.org/10.2307/2580062>
- Loukil, N., & Yousfi, O. (2016). Does gender diversity on corporate boards increase risk-taking? *Canadian Journal of Administrative Sciences*, 33(1), 66–81. <https://doi.org/10.1002/cjas.1326>
- Lynch, S. M., & Brown, J. S. (2011). Stratification and inequality over the life course. In R. H. Binstock & L. K. George (Eds.), *Handbook of aging and the social sciences* (pp. 105–117). Academic Press. <https://doi.org/10.1016/B978-0-12-380880-6.00008-3>
- Mangematin, V., Lemarié, S., Boissin, J.-P., Catherine, D., Corolleur, F., Coronini, R., & Trommetter, M. (2003). Development of SMEs and heterogeneity of trajectories: The case of biotechnology in France. *Research Policy*, 32(4), 621–638. [https://doi.org/10.1016/s0048-7333\(02\)00045-8](https://doi.org/10.1016/s0048-7333(02)00045-8)
- Marco, R. (2012). Gender and economic performance: Evidence from the Spanish hotel industry. *International Journal of Hospitality Management*, 31(3), 981–989. <https://doi.org/10.1016/j.ijhm.2011.12.002>
- Martín-Ugedo, J. F., Mínguez-Vera, A., & Rossi, F. (2019). Female directors and firm performance in Italian and Spanish listed firms. *Academia Revista Latinoamericana de Administración*, 32(3), 411–436. <https://doi.org/10.1108/arla-06-2018-0124>
- Martín-Ugedo, J. F., Mínguez-Vera, A., & Palma-Martos, L. (2018). Female CEOs, returns and risk in Spanish publishing firms. *European Management Review*, 15(1), 111–120. <https://doi.org/10.1111/emre.12132>
- Maydeu-Olivares, A., Shi, D., & Rosseel, Y. (2019). Instrumental variables two-stage least squares (2SLS) vs. maximum likelihood structural equation modeling of causal effects in linear regression models. *Structural Equation Modeling A Multidisciplinary Journal*, 26(6), 876–892. <https://doi.org/10.1080/10705511.2019.1607740>

- McGuinness, P. B., Vieito, J. P., & Wang, M. (2017). The role of board gender and foreign ownership in the CSR performance of Chinese listed firms. *Journal of Corporate Finance*, 42, 75–99. <https://doi.org/10.1016/j.jcorpfin.2016.11.001>
- Menicucci, E., Paolucci, G., & Paoloni, N. (2019). Does gender matter for hotel performance? Evidence from the Italian hospitality industry. *International Journal of Tourism Research*, 21(5), 625–638. <https://doi.org/10.1002/jtr.2286>
- Metcalf, B. D. (2007). Gender and human resource management in the Middle East. *International Journal of Human Resource Management*, 18(1), 54–74. <https://doi.org/10.1080/09585190601068292>
- Mkhetwa, M., & Msweli, P. (2011). The impact of female business leaders on the performance of listed companies in South Africa. *South African Journal of Economic and Management Sciences*, 14, 1–7.
- Mohan, N. (2014). A review of the gender effect on pay, corporate performance and entry into top management. *International Review of Economics & Finance*, 34, 41–51. <https://doi.org/10.1016/j.iref.2014.06.005>
- Moore, D. P. (1990). An examination of present research on the female entrepreneur? Suggested research strategies for the 1990's. *Journal of Business Ethics*, 9(4-5), 275–281. <https://doi.org/10.1007/bf00380327>
- Moreno-Gómez, J., & Calleja-Blanco, J. (2018). The relationship between women's presence in corporate positions and firm performance: The case of Colombia. *International Journal of Gender and Entrepreneurship*, 10(1), 83–100. <https://doi.org/10.1108/ijge-10-2017-0071>
- Moreno-Gómez, J., Lafuente, E., & Vaillant, Y. (2018). Gender diversity in the board, women's leadership and business performance. *Gender in Management An International Journal*, 33(2), 104–122. <https://doi.org/10.1108/gm-05-2017-0058>
- Nakamura, A., & Nakamura, M. (1998). Model specification and endogeneity. *Journal of Econometrics*, 83(1-2), 213–237. [https://doi.org/10.1016/s0304-4076\(97\)00070-5](https://doi.org/10.1016/s0304-4076(97)00070-5)
- Nekhili, M., Chakroun, H., & Chtioui, T. (2018). Women's leadership and firm performance: Family versus nonfamily firms. *Journal of Business Ethics*, 153(2), 291–316. <https://doi.org/10.1007/s10551-016-3340-2>
- Nicholls-Nixon, C. L., Davila Castilla, J. A., Sanchez Garcia, J., & Rivera Pesquera, M. (2011). Latin America management research: Review, synthesis, and extension. *Journal of Management*, 37(4), 1178–1227. <https://doi.org/10.1177/0149206311403151>
- Nili, Y. (2019). Beyond the numbers: Substantive gender diversity in boardrooms. *Indiana Law Journal*, 94(1), 145–202. <https://doi.org/10.2139/ssrn.3117131>
- Noland, M., Moran, T., & Kotschwar, B. R. (2016). Is gender diversity profitable? Evidence from a global survey. *SSRN Electronic Journal*. 1–35. <https://doi.org/10.2139/ssrn.2729348>
- Ntim, C. G. (2015). Board diversity and organizational valuation: Unravelling the effects of ethnicity and gender. *Journal of Management & Governance*, 19(1), 167–195. <https://doi.org/10.1007/s10997-013-9283-4>
- Perez-Batres, L. A., Pisani, M. J., & Doh, J. P. (2012). An assessment of the role of Latin America in the core international business literature (2001–2010). *Latin American Business Review*, 13(4), 263–287. <https://doi.org/10.1080/10978526.2012.749076>
- Perryman, A. A., Fernando, G. D., & Tripathy, A. (2016). Do gender differences persist? An examination of gender diversity on firm performance, risk, and executive compensation. *Journal of Business Research*, 69(2), 579–586. <https://doi.org/10.1016/j.jbusres.2015.05.013>
- Porcena, Y.-R., Parboteeah, K. P., & Mero, N. P. (2021). Diversity and firm performance: Role of corporate ethics. *Management Decision*, 59, 2620–2644. ahead-of-p(ahead-of-print). <https://doi.org/10.1108/md-01-2019-0142>
- Post, C., & Byron, K. (2015). Women on boards and firm financial performance: A meta-analysis. *Academy of Management Journal*, 58(5), 1546–1571. <https://doi.org/10.5465/amj.2013.0319>
- Ramsey, L. R. (2017). Agentic traits are associated with success in science more than communal traits. *Personality and Individual Differences*, 106, 6–9. <https://doi.org/10.1016/j.paid.2016.10.017>
- Recio, A. (1997). *Trabajo, personas, mercados: manual de economía laboral*. [https://books.google.cl/books?hl=es&lr=&id=FqckDVm9-8kC&oi=fnd&pg=PA9&dq=Recio,+A.++\(1997\),+Trabajo,+personas,+mercados,+Barcelona,+Icaria.+Capítulo+10&ots=hkwpXW0PyH&sig=qrVJ4ZdmtDhDaJYjpxR3zcYFGFU](https://books.google.cl/books?hl=es&lr=&id=FqckDVm9-8kC&oi=fnd&pg=PA9&dq=Recio,+A.++(1997),+Trabajo,+personas,+mercados,+Barcelona,+Icaria.+Capítulo+10&ots=hkwpXW0PyH&sig=qrVJ4ZdmtDhDaJYjpxR3zcYFGFU)
- Reddy, S., & Jadhav, A. M. (2019). Gender diversity in boardrooms—A literature review. *Cogent Economics & Finance*, 7(1), 1644703. <https://doi.org/10.1080/23322039.2019.1644703>
- Reinert, R. M., Weigert, F., & Winnefeld, C. H. (2016). Does female management influence firm performance? Evidence from Luxembourg banks. *Financial Markets and Portfolio Management*, 30(2), 113–136. <https://doi.org/10.1007/s11408-016-0266-8>
- Rodríguez-Domínguez, L., García-Sánchez, I.-M., & Gallego-Álvarez, I. (2012). Explanatory factors of the relationship between gender diversity and corporate performance. *European Journal of Law and Economics*, 33(3), 603–620. <https://doi.org/10.1007/s10657-010-9144-4>
- Rose, C. (2007). Does female board representation influence firm performance? The Danish evidence. *Corporate Governance*, 15(2), 404–413. <https://doi.org/10.1111/j.1467-8683.2007.00570.x>
- Rucker, D. D., Galinsky, A. D., & Magee, J. C. (2018). The agentic–communal model of advantage and disadvantage: how inequality produces similarities in the psychology of power, social class, gender, and race. In J. M. Olson (Eds.), *Advances in experimental social psychology* (pp. 71–125). Academic Press. <https://doi.org/10.1016/bs.aesp.2018.04.001>
- Sargan, J. D. (1958). The estimation of economic relationships using instrumental variables. *Econometrica*, 26(3), 393. <https://doi.org/10.2307/1907619>
- Sargan, J. D. (1988). Testing for misspecification after estimating using instrumental variables. *Contributions to Econometrics: John Denis Sargan*, 1, 213–235.
- Sarrió, M., Barberá, E., Ramos, A., & Candela, C. (2002). El techo de cristal en la promoción profesional de las mujeres. *Revista de Psicología Social: International Journal of Social Psychology*, 17(2), 167–182. <https://doi.org/10.1174/021347402320007582>

- Schmidt, P. (2020). *Econometrics (first)*. CRC Press.
- SCImago, & Scopus. (2020). *Scimago Journal & Country Rank. Journal Rankings*. <https://www.scimagojr.com/>
- Simonsen, M., & Skipper, L. (2006). The costs of motherhood: An analysis using matching estimators. *Journal of Applied Economics*, 21(7), 919–934. <https://doi.org/10.1002/jae.893>
- Singh, A., Singhanian, S., & Sardana, V. (2019). Do women on boards affect Firm's financial performance? Evidence from Indian IPO firms. *Australasian Accounting Business and Finance Journal*, 13(2), 53–68. <https://doi.org/10.14453/aabfj.v13i2.4>
- Singhathep, T., & Pholphirul, P. (2015). Female CEOs, firm performance, and firm development: Evidence from Thai manufacturers. *Gender Technology and Development*, 19(3), 320–345. <https://doi.org/10.1177/0971852415596865>
- Spence, J. T., & Helmreich, R. L. (1979). Comparison of masculine and feminine personality attributes and sex-role attitudes across age groups. *Developmental Psychology*, 15(5), 583–584. <https://doi.org/10.1037/h0078091>
- Stock, J., & Watson, M. (2012). *Introducción a la Econometría (3ª)*. Pearson Educación. <https://doi.org/10.15446/ideasyvalores.v68n171.79906>
- Strydom, M., Au Yong, H. H., & Rankin, M. (2017). A few good (wo)men? Gender diversity on Australian boards. *Australian Journal of Management*, 42(3), 404–427. <https://doi.org/10.1177/0312896216657579>
- Strøm, R. Ø., D'Espallier, B., & Mersland, R. (2014). Female leadership, performance, and governance in microfinance institutions. *Journal of Banking & Finance*, 42, 60–75. <https://doi.org/10.1016/j.jbankfin.2014.01.014>
- Sullivan, D. M., & Meek, W. R. (2012). Gender and entrepreneurship: A review and process model. *Journal of Managerial Psychology*, 27(5), 428–458. <https://doi.org/10.1108/02683941211235373>
- Terjesen, S., Couto, E. B., & Francisco, P. M. (2016). Does the presence of independent and female directors impact firm performance? A multi-country study of board diversity. *Journal of Management & Governance*, 20(3), 447–483. <https://doi.org/10.1007/s10997-014-9307-8>
- Terjesen, S., Sealy, R., & Singh, V. (2009). Women Directors on Corporate Boards: A Review and Research Agenda. *Corporate Governance*, 17(3), 320–337. <https://doi.org/10.1111/j.1467-8683.2009.00742.x>
- Titus, M. A. (2007). Detecting selection bias, using propensity score matching, and estimating treatment effects: An application to the private returns to a master's degree. *Research in Higher Education*, 48(4), 487–521.
- Tondji, J. (2022). Overconfidence and welfare in a differentiated duopoly. *Managerial and Decision Economics*, 43(3), 751–767. <https://doi.org/10.1002/mde.3416>
- Ullah, I., Fang, H., & Jebran, K. (2019). Do gender diversity and CEO gender enhance firm's value? Evidence from an emerging economy. *Corporate Governance*, 20(1), 44–66. <https://doi.org/10.1108/cg-03-2019-0085>
- Unite, A. A., Sullivan, M. J., & Shi, A. A. (2019). Board diversity and performance of philippine firms: Do women matter? *International Advances in Economic Research*, 25(1), 65–78. <https://doi.org/10.1007/s11294-018-09718-z>
- Valls Martínez, M. del C., & Cruz Rambaud, S. (2019). Women on corporate boards and firm's financial performance. *Women's Studies International Forum*, 76, 102251. <https://doi.org/10.1016/j.wsif.2019.102251>
- Vinha, K. (2006). A Primer on Propensity Score Matching Estimators. In *CEDE 2006-13* (Vol. 7191, pp. 1–28). Universidad de los Andes.
- Vu, T.-H., Nguyen, V.-D., Ho, M.-T., & Vuong, Q.-H. (2019). Determinants of Vietnamese listed firm performance: Competition, Wage, CEO, firm size, age, and International Trade. *Journal of Risk and Financial Management*, 12(2), 62–19. <https://doi.org/10.3390/jrfm12020062>
- Wang, G., Holmes, R. M., Devine, R. A., & Bishoff, J. (2018). CEO gender differences in careers and the moderating role of country culture: A meta-analytic investigation. *Organizational Behavior and Human Decision Processes*, 148, 30–53. <https://doi.org/10.1016/j.obhdp.2018.04.002>
- Wang, Y., & Bellemare, M. F. (2019). *Lagged variables as instruments*. Working Paper, Department of Applied Economics, University of Minnesota.
- Watson, J. (2002). Comparing the performance of male-and female-controlled businesses: Relating outputs to inputs. *Entrepreneurship Theory and Practice*, 26(3), 91–100. <https://doi.org/10.1177/104225870202600306>
- Watson, J. (2020). Exposing/correcting SME underperformance myths. *International Journal of Gender and Entrepreneurship*, 12(1), 77–88. <https://doi.org/10.1108/ijge-04-2019-0086>
- Watson, J., Stuetzer, M., & Zolin, R. (2017). Female underperformance or goal orientated behavior? *International Journal of Gender and Entrepreneurship*, 9(4), 298–318. <https://doi.org/10.1108/ijge-03-2017-0015>
- Westhead, P. (2003). Comparing the performance of male- and female-controlled businesses. *Journal of Small Business and Enterprise Development*, 10(2), 217–224. <https://doi.org/10.1108/14626000310473265>
- Wille, B., Wiernik, B. M., Vergauwe, J., Vrijdags, A., & Trbovic, N. (2018). Personality characteristics of male and female executives: Distinct pathways to success? *Journal of Vocational Behavior*, 106, 220–235. <https://doi.org/10.1016/j.jvb.2018.02.005>
- Wooldridge, J. (Ed.). (2010). *Econometric Analysis of cross section and panel data* (2nd ed.). MIT Press.
- Wooldridge, J. (2020). *Introductory econometrics: A modern approach (seven)*. Cengage.
- Wu, D.-M. (1974). Alternative tests of independence between stochastic regressors and disturbances: Finite sample results. *Econometrica*, 42(3), 529. <https://doi.org/10.2307/1911789>
- Yogo, M. (2004). Estimating the elasticity of intertemporal substitution when instruments are weak. *Review of Economics and Statistics*, 86(3), 797–810. <https://doi.org/10.1162/0034653041811770>
- Zolin, R., Stuetzer, M., & Watson, J. (2013). Challenging the female underperformance hypothesis. *International Journal of Gender and Entrepreneurship*, 5(2), 116–129. <https://doi.org/10.1108/17566261311328819>