Data in Brief 65 (2026) 112488

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Dataset on Spanish medium-sized family firms: Linking socioemotional wealth, HRM practices, and financial indicators

J. Samuel Baixauli-Soler[a], María Belda-Ruiz[a], Gabriel Lozano-Reina[a,*], Juan David Peláez-León[b], Gregorio Sánchez-Marín[c]

[a] Department of Management and Finance, Faculty of Economics and Business, University of Murcia, Spain
[b] Department of Management and Organization, Universidad del Valle, Colombia
[c] Department of Economics and Business, Faculty of Economics, Business and Tourism, University of Alcalá, Spain

ARTICLE INFO

Article history:
Received 1 October 2025
Revised 19 December 2025
Accepted 14 January 2026
Available online 19 January 2026

Dataset link: Family firms: SEW, HRM practices, and financial indicators (Original data)

Keywords:
Family firms
Socioemotional wealth (SEW)
Human resource management (HRM)
Survey data
Financial indicators
SABI database
Spain

ABSTRACT

This article presents a dataset on 508 medium-sized Spanish family firms, collected between March and June 2016 through structured telephone interviews with CEOs or HR directors. The questionnaire covered four dimensions: family involvement and socioemotional wealth (SEW), human resource management (HRM) practices, financial strategies, and managerial demographics. To complement survey data, financial indicators were extracted from the SABI (*Sistema de Análisis de Balances Ibéricos*) database. The dataset integrates subjective managerial assessments with objective firm-level information, offering a unique resource for research on family business management, HRM, and financial policies. Variables include firm ownership and management, generational structures, SEW priorities, human capital and HRM practices, financial goals and capital access, as well as managers' demographic characteristics. The database is released in cleaned, anonymized, and fully documented form (together with the questionnaire and a detailed codebook), enabling replication,

* Corresponding author.
E-mail address: gabriel.lozano@um.es (G. Lozano-Reina).
Social media: @Samuelbaixauli (J.S. Baixauli-Soler), @MariaBeldaRuiz (M. Belda-Ruiz), @LozanoReina (G. Lozano-Reina), @juanda174 (J.D. Peláez-León)

https://doi.org/10.1016/j.dib.2026.112488

comparative studies, and meta-analyses on family firms and related organizational topics.

Specifications Table

Subject	Social Sciences
Specific subject area	Family business, human resource management (HRM), and financial strategies of Spanish family firms
Type of data	Raw, cleaned, labelled, anonymized
Data collection	Data were collected through structured telephone interviews with CEOs or HR directors of 508 Spanish medium-sized family firms (March-June 2016). The questionnaire included validated scales (e.g., SEW, HRM practices). Stratified random sampling was applied by sector and firm age. Financial and structural indicators were retrieved from the SABI (*Sistema de Análisis de Balances Ibéricos*) database (2015 financial year). All data were checked, cleaned, and anonymized.
Data source location	• Country: Spain (nationwide, industry and services) • Primary data source: Structured managerial survey (2016) • Secondary data source: SABI database (2015)
Data accessibility	• Repository name: Zenodo • Data identification number: 10.5281/zenodo.17723051 • Direct URL to data: https://doi.org/10.5281/zenodo.17225495
Related research article	Not applicable.

1. Value of the Data

- These data provide unique firm-level information on 508 Spanish medium-sized family firms, combining a structured managerial survey with objective financial indicators from the SABI (Sistema de Análisis de Balances Ibéricos) database. This dual-source structure offers a robust empirical basis for secondary analyses.
- The dataset enables multiple operationalizations of family involvement and socioemotional wealth (SEW), going beyond binary definitions of family firms. These measures support alternative classifications, cross-study comparisons, and meta-analyses on family firm heterogeneity.
- Comprehensive indicators of human resource management (HRM) practices are included, based on the ability–motivation–opportunity (AMO) model, which includes selection, training, appraisal, compensation, and participation practices. These measures allow cross-sectional benchmarking and replication of HRM-related findings in family business and organizational research.
- The integration of subjective managerial assessments with objective financial indicators enhances opportunities for robustness checks. Scholars can re-use the data to validate perceptions against financial outcomes or to explore new questions in governance, HRM, and family business research.
- The dataset is released together with full supporting documentation –including the questionnaire and a detailed codebook that explains the stratification design and the construction of the post-stratification weight variable– thereby ensuring transparency, reproducibility, and efficient reuse.
- Compared with previous Data in Brief datasets on family firms, such as Arz & Kuckertz's [1] or Meroño-Cerdan et al.'s [2], this dataset offers a broader and more integrated perspective. It jointly captures SEW, detailed HRM practices, employee outcomes, financial strategies, and objective SABI-based indicators for a stratified sample of medium-sized family firms, enabling

analyses of how SEW shapes HRM and financial decisions –an area rarely covered in prior datasets.

2. Background

Family firms constitute the dominant organizational form worldwide, where they account for more than 80% of companies and generate a substantial share of GDP and employment [3]. Their governance and management are influenced not only by financial objectives but also by family-related goals such as control, reputation, and transgenerational continuity [4]. These non-economic objectives are commonly conceptualized as socioemotional wealth (SEW), which has become a central framework for understanding family firm behavior [5,6].

The dataset was compiled to address the scarcity of large-scale data that simultaneously captures family ownership and governance, SEW, human resource management (HRM) practices, financial strategies, and demographics. A structured questionnaire was designed and administered to CEOs or HR directors of Spanish medium-sized family firms. The instrument drew on validated scales in the literature, including the SEWi scale [7], which has been further validated in subsequent research [8,9]. For HRM practices, we relied on established measures grounded in the ability–motivation–opportunity (AMO) model, whose validity has also been supported in recent empirical studies [10,11]. Data collection took place between March and June 2016 through structured telephone interviews, complemented with financial and structural information extracted from the SABI (*Sistema de Análisis de Balances Ibéricos*) database for the fiscal year 2015[1].

This article makes the dataset and its documentation (codebook and questionnaire) publicly available to ensure transparency and facilitate reuse in family business, HRM, and governance research.

3. Data Description

The dataset is organized into four main blocks corresponding to the questionnaire given to CEOs or HR directors of Spanish family firms, complemented with economic-financial variables extracted from the SABI database. In total, it contains information for 508 firms, with survey data aligned with financial and structural indicators. Several procedural measures were used to reduce perceptual and social-desirability biases. Importantly, the dataset includes a variable identifying whether the respondent (CEO or HR director) belongs to the owning family, which is essential for assessing the potential impact of respondent characteristics on data reliability.

3.1. Block I. Family/ownership dimension

This block allows both the structural involvement of the family (ownership, management, succession) and the socioemotional priorities guiding family firms to be quantified. It relies exclusively on data collected through the questionnaire. Specifically, it includes:

- Self-identification as a family firm (*P1_def*): dichotomous variable coded as 1 = family firm, 0 = non-family firm. This question served as a screening item. If the response was “No”, the survey was discontinued and the firm was excluded from the final sample.
- Ownership distribution (*P2_pro*): percentage of shares held by (a) family owners, (b) non-family owners, and (c) institutional shareholders (banks, pension funds, mutual funds).

[1] Although the survey was conducted between March and June 2016, most questionnaire items referred retrospectively to the previous year (2015) or earlier (e.g., investment in employee training hours). For consistency, the financial and structural indicators extracted from the SABI database correspond to fiscal year 2015.

- Family generation leading the company (*P3_gen*): categorical variable coded as 1 = first generation (founding), 2 = second generation, 3 = third generation, 4 = fourth generation or later.
- Managerial positions: total number of managerial positions in the firm (*P4a_dire*) and the number of those positions held by family members (*P4b_dirf*).
- Socioemotional wealth (SEW) (*P5_1_sew* to *P5_13_sew*): measured with 13 items on a 5-point Likert scale (where 1 = not important, and 5 = very important), adapted from Berrone et al. [5] and Debicki et al. [7]. More specifically, items capture the importance attached to preserving family unity, succession to the next generation, maintaining the family dynasty, sustaining family values, protecting family reputation, community recognition, relationships with stakeholders, building social capital, treating non-family employees as family, family members' well-being and lifestyle, family harmony, consideration of family needs in decision-making, and ensuring the happiness of non-employed family members.

3.2. Block II. HRM dimension

This block focuses on HRM practices as well as employee drivers and outcomes. It also relies exclusively on data collected through the questionnaire. Specifically, it includes:

- Training activities: number of non-managerial employees trained during 2015 (*P6a_efor*), and total number of training hours delivered (*P6b_hfor*).
- Sector-specific key non-managerial employee category (*P7_sect*): identification of the most important category of non-managerial employees for the company's core activity, distinguishing between service (salespeople, customer service, operators, others) and industrial firms (designers, engineers, maintenance staff, operators, others).
- HRM practices (*P8_1_HR* to *P8_18_HR*): measured through 18 items on a 5-point Likert scale (where 1 = total disagreement, and 5 = total agreement), adapted from Jiang et al. [12]. These items involved high-performance HRM practices around the AMO model [13]. HRM practices include recruitment and selection (e.g., adequacy of candidate, skills-based or requirements-based selection), training and development (e.g., continuous training, investment in training, quality-oriented training, broad training beyond skills; career plans), performance appraisal (e.g., evaluation with quantifiable objectives, multi-source evaluation, feedback), compensation (e.g., pay-for-performance, equity in compensation, incentives), and participation and work-life balance (e.g., suggestions for improvement, participation in decision-making, concern for conciliation, adapted schedules).
- Factors influencing training decisions (*P9_1_for* to *P9_9_for*): measured with nine items on a 5-point Likert scale (where 1 = minimal influence, and 5 = maximum influence). Specifically, items include compliance with legal regulations, availability of subsidized training, union/works council influence, productivity needs, HR strategy, recommendations from business associations, quality certifications, benchmarking with successful companies, and sectoral norms.
- Employees' affective commitment (*P10_1_co* to *P10_3_co*): measured with three items on a 5-point Likert scale (where 1 = total disagreement, and 5 = total agreement), capturing the extent to which employees consider the company to be personally significant, how emotionally attached they feel, and how much they see themselves as part of a family.
- Human capital (*P11_1_ch* to *P11_5_ch*): assessed via five items on a 5-point Likert scale (where 1 = total disagreement, and 5 = total agreement), measuring whether employees are perceived as highly qualified, the best in the sector, creative, experts in their functions, and innovative (developing new ideas and knowledge).
- Employee performance (*P12_1_dl* to *P12_5_dl*): measured with five items on a 5-point Likert scale (where 1 = total disagreement, and 5 = total agreement), including doing an excellent job, providing added value, few customer complaints, resource efficiency, and overall exceptional performance.

- Workforce composition: number of non-managerial employees (*P13a_emp*) and how many of them belong to the owning family (*P13b_emp*).

This block not only reports HRM practices formally adopted by firms but also captures their perceived drivers and outcomes (commitment, human capital, and job performance), therefore providing a comprehensive picture of the management of human resources in family firms.

3.3. Block III. Financial dimension

This block explores firms' financial goals, access to external financing, and the allocation of raised funds. Specifically, it includes the following variables obtained from the questionnaire:

- Financial objectives (*P14_finz*): measured on a 5-point scale ranging from 1 = maximum stability/solvency to 5 = maximum profitability, capturing the relative priority assigned by the firm to stability versus profitability in its financial objectives.
- Access to external funding (*P15_1, P15_2, P15_3*): binary variables (1 = yes, 0 = no) indicating whether the firm attempted to raise funds in the past three years through (a) financial loans, (b) capital (equity funds), or (c) other sources of financing.
- Success in funding attempts: when applicable, firms rated their success on a 5-point scale (1 = total failure, 5 = total success).
- Destination of funds (*P16_1* to *P16_7*): binary indicators (1 = yes, 0 = no) specifying whether the funds raised were allocated to: keeping business operations, increasing domestic activities, financing exports, financing innovation and R&D, carrying out mergers or acquisitions, developing international activities, or other purposes.

These indicators capture the financial orientation of firms, their ability to secure external resources, and the allocation of funding to strategic uses. Moreover, survey responses are complemented with objective financial and structural indicators obtained from the SABI database for the fiscal year 2015, including:

- Firm size and structure: number of employees (*B1_empl*), operating revenues (*B2_inc*, in thousands of Euros), and total assets (*B3_assets*, in thousands of Euros).
- Profitability and performance indicators: earnings before interest, taxes, depreciation, and amortization (EBITDA) (*B4_ebitda*), shareholders' equity (*B5_equity*), current liabilities (*B6_liabil*), current assets (*B7_assetscur*), return on assets (*B8_ROA*), and return on equity (*B9_ROE*).

This block allows the comparison between subjective financial strategies and objective performance indicators, enabling a more robust analysis of the financial behavior of family firms.

3.4. Block IV. Demographic dimension

This last block provides contextual and demographic information at both the firm and managerial levels. Managerial-level variables are obtained from the questionnaire (CEO and HR director) and include the following:

- Respondent's managerial role (*P17_1*): binary indicator identifying whether the respondent holds a key managerial role in the firm (CEO and/or HR director), coded as 1 = yes and 0 = no.
- Tenure (*P17_2*): number of years the CEO or HR director has worked in the firm.
- Age (*P17_3*): CEO's or HR director's age, self-reported in years.
- Gender (*P17_4*): CEO's or HR director's gender, coded as 1 = female, 0 = male.
- Educational level (*P17_5*): CEO's or HR director's educational level, coded as 1 = university, 2 = secondary-vocational training or high school, 0 = primary.

Table 1
Measurement quality of the multi-item constructs.

Construct	Items	Cronbach's α	% Variance	Range of factor loadings
Socioemotional wealth (SEW)	13	0.919	52.03%	0.52 – 0.85
HRM practices	18	0.924	44.82%	0.55 – 0.75
Training drivers	9	0.812	40.54%	0.56 – 0.75
Affective commitment	3	0.847	77.27%	0.85 – 0.91
Human capital	5	0.868	65.83%	0.78 – 0.84
Employee performance	5	0.872	66.70%	0.74 – 0.87

- Family membership (*P17_6*): dichotomous variable capturing whether the CEO or HR director belongs to the owning family (1 = yes, 0 = no).

Firm-level variables are obtained from the SABI database and include the following:

- Firm age, which is measured in two ways. First, as a continuous variable (in years) (*B10a_age*), calculated from the incorporation date up to March 2016 (reference date when the survey began). Second, as a categorical variable (*B10b_agecod*) coded as 1 = between 0 and 10 years, 2 = more than 10 and less than 30 years, 3 = more than 30 and less than 50 years, and 4 = more than 50 years.
- Industry classification (*B11_sect*): dichotomous variable distinguishing services (value 1) and industry (value 0).
- Legal form (*B12_legalform*): coded as 1 = limited liability company, 0 = public limited company.
- Stock exchange listing (*B13_stockexch*): coded as 1 = currently listed, 2 = formerly listed, 0 = not listed.

3.5. Sampling adjustments

The dataset also includes a post-stratification weight variable (*weight_ps*), computed to correct minor deviations between the population and the achieved sample across the sector × age strata. The computation method and recommended usage are fully documented in the accompanying codebook.

3.6. Assessment of measurement properties

To document the internal consistency of the constructs included in the dataset, reliability analyses and principal component analyses were conducted. Results are summarized in Table 1. Cronbach's alpha coefficients range from 0.81 to 0.92 across scales. For each construct, the first principal component accounts for 41% to 77% of the variance, with item loadings exceeding 0.50, indicating satisfactory internal coherence. It should be noted that both the SEW and HRM constructs are multidimensional, and more detailed analyses of their subdimensions may refer to the validated frameworks available in previous studies (e.g. [10,9,11]).

4. Experimental Design, Materials and Methods

4.1. Sampling frame and fieldwork design

The sampling frame was constructed from an extensive database developed by the Family Business Firms Institute in Spain [14] using firm-level information from the Spanish SABI database. This source systematically identifies medium-sized family firms that meet formal family governance and ownership criteria. Specifically, firms in which (i) at least one family mem-

Table 2
Population, theoretical allocation, and final achieved sample.

Sector	Age (years)	Population (n)	Population (%)	Theoretical allocation (n)	Final sample (n)	Post-stratification weights
Industry	0-10	53	4.60%	7	5	1.3737
	10-30	527	45.40%	67	69	0.9898
	30-50	477	41.10%	61	60	1.0303
	>50	103	8.90%	13	13	1.0268
Total industry		**1,160**	**29.60%**	**148**	**147**	
Services	0-10	234	8.50%	30	30	1.0108
	10-30	1,783	64.60%	227	218	1.0599
	30-50	630	22.80%	80	99	0.8247
	>50	113	4.10%	14	14	1.0460
Total services		**2,760**	**70.40%**	**352**	**365**	
Total		**3,920**	**100%**	**500**	**512**	

Note: Of the 512 interviews achieved, four cases were excluded during data cleaning, thus yielding a final valid sample of 508 firms.

ber participates in governance or management, and (ii) the family holds a significant ownership stake (either one member with ≥5% or several members jointly with ≥20%). Although this database does provide reliable coverage of established and medium-sized family firms, it does not capture the full universe of Spanish family firms (particularly very small, informal, or unregistered businesses for which no national registry exists). As is common in family business research –where the true population is inherently unknown– the resulting sampling frame is representative of this well-defined segment rather than of all Spanish family firms. The final target population consisted of 3,920 firms, which were used to define the sampling strata by sector and firm age, as shown in Table 2.

A stratified random sampling procedure with proportional allocation was implemented. Strata were defined by the combination of sector (industry vs. services) and firm age (<10 years, 10-30 years, 30-50 years, >50 years). Quotas were monitored throughout the fieldwork to ensure that the achieved sample approximated the proportional distribution of the population. A theoretical sample size of 500 firms was determined; for a finite population (N=3,920), at a 95% confidence level and assuming p=q=0.5, the maximum sampling error was ±4.06%. The final valid dataset comprises 512 firms, exceeding the planned target.

Although proportional allocation was followed, minor deviations did emerge during fieldwork (e.g., services firms aged 30-50 were over-represented, whereas those aged 10-30 were slightly under-represented). In addition, detailed call outcomes (refusals, non-contacts, ineligible cases) were not recorded, thus preventing a full non-response analysis. In order to address these issues and enhance representativeness relative to the sampling frame, we computed post-stratification weights (*weight_ps*) based on population margins across the sector × age strata using:

$$weight\_ps = \frac{(N_h/N)}{(n_h/n)}$$

where N_h and n_h denote the population and sample counts in stratum h, and N and n denote the total population and sample sizes, respectively. The variable *weight_ps* is included in the dataset and documented in the codebook.

Structured telephone interviews were conducted with the CEO or HR director of each firm. Fieldwork was carried out between March and June 2016 by trained interviewers from *CSA Consultores*, a professional market research firm. Interviewers followed standardized scripts and fixed item ordering, and multiple call attempts were made to contact each selected firm. Interviews lasted approximately 20-30 minutes, depending on respondent availability and firm complexity. Respondents were offered a summary research report as a participation incentive.

The initial dataset consisted of 512 interviews. During data cleaning, four cases were excluded due to incomplete or inconsistent responses, resulting in a final valid sample of 508 firms. The effective response rate was 12.96%, calculated as 508 valid interviews divided by the 3,920 eligible firms in the sampling frame. This rate aligns with findings in the family business survey literature, where response rates are structurally lower than in general management research due to gatekeeping, lack of incentives, and the seniority of the target respondents. In this regard, a meta-analysis of 126 family business survey studies reports an average response rate of around 21% [15], and recent empirical work documents rates near 17.8% as comparable benchmarks in this domain [16]. Accordingly, our achieved rate falls within the lower range commonly observed among family firm surveys.

4.2. Survey instrument

The questionnaire instrument (see supplementary file) was designed to capture the main dimensions of family firm behavior and strategy. It consisted of four thematic blocks: (i) family/ownership, (ii) HRM, (iii) financial strategies, and (iv) demographics. The questionnaire included closed-ended questions (dichotomous, categorical, and numeric) and multi-item Likert-type scales. All items were coded, labelled, and standardized in the database.

The instrument referred to multiple temporal frames, which implies different recall windows depending on the type of item. Questions relating to activities performed during the 2015 fiscal year (e.g., number of employees trained, total training hours) imply a recall window of approximately 3 to 15 months depending on the interview date (March-June 2016). Items referring to the "last three years" (e.g., SEW orientation, HRM practices, access to external financing, use of funds) have a fixed reference period by design. To reduce recall burden, interviewers used explicit anchoring cues embedded in the questionnaire, such as "during 2015…", "in the last three years…", or "compared with your competitors…". Recall errors may be more pronounced for training hours and HRM practices, although the latter tend to be structurally stable over time.

Several procedural remedies were implemented to reduce perceptual and social-desirability biases. Interviewers used standardized scripts with neutral wording and were instructed to avoid evaluative feedback or leading formulations. The questionnaire was structured into conceptually distinct blocks (SEW, HRM practices, financial decisions, demographics), with fixed ordering to minimize priming and contamination across constructs. Screening ensured that respondents were either the CEO or the HR director; i.e., individuals expected to possess an accurate knowledge of the topics addressed. In addition, post-survey validation was performed by cross-checking perceptual responses with objective SABI indicators.

Interviews were conducted through a Computer-Assisted Telephone Interviewing (CATI) platform, which enforced fixed item ordering, automatic routing, and real-time consistency checks. As mentioned above, data collection was performed by *CSA Consultores*. All interviews were carried out by experienced fieldworkers employed by the company. Prior to fieldwork, interviewers participated in a dedicated training session jointly delivered by the project's principal investigators and by CSA supervisors. Training covered standardized scripts, contextual clarifications for key constructs, neutral probing techniques, and procedures designed to minimize interviewer effects. CSA supervisors monitored call quality and adherence to the protocol throughout the field period so as to ensure consistency.

To guarantee construct validity, items were drawn from previously validated scales and adapted to the Spanish context. SEW was measured with a 13-item scale based on Berrone et al. [5] and Debicki et al. [7], which has shown high internal consistency in prior studies (e.g., $\alpha = 0.913$; [8]). HRM practices were assessed using multi-item Likert scales grounded in the AMO framework [13,12], with reported reliability levels between 0.84 and 0.90 [10]. Financial strategy questions followed standard operationalizations of funding access, allocation, and priorities. All items were reviewed and linguistically adapted to ensure cultural relevance in Spain (e.g., references to works councils and training subsidies). The questionnaire was pre-tested with

a small group of CEOs and HR directors to evaluate clarity, sequence, and relevance. Feedback resulted in minor wording adjustments and improvements in item flow. The final instrument was approved before fieldwork.

4.3. Data integration and cleaning

Survey responses were matched with objective economic and structural indicators extracted from the SABI database for the fiscal year 2015. Matching was carried out using firm identifiers, thereby allowing subjective managerial assessments to be compared directly with objective firm-level information. This integration provides a dual perspective: perceptual evaluations of CEOs or HR directors, and audited firm-level financial records.

A multi-step procedure was applied to verify data integrity before finalizing the dataset. First, survey responses were validated to ensure logical coherence across related variables; specifically, ownership distribution percentages (*P2a-P2c*) had to come to 100%; the number of family directors (*P4b_dirf*) could not exceed the total number of directors (*P4a_dire*); and HR-related counts (e.g., number of employees trained, training hours) were reviewed against firm size. Self-identification as a family firm was also cross-checked against objective indicators of family ownership and involvement in governance to ensure consistency with the criteria used in the sampling frame. Any inconsistencies were flagged for later review during the cleaning process.

Second, responses outside allowable ranges (e.g., Likert scales coded 1-5) were flagged. Implausible values (e.g., training hours exceeding feasible ranges) were reviewed and corrected if they were clearly input errors, or they were recoded as missing. Third, survey responses were cross-validated with SABI data. In the event of discrepancy, SABI values were retained as the authoritative source. Fourth, from the 512 interviews obtained, two were discarded due to incomplete responses (substantial missing data in key blocks), and two due to inconsistencies that could not be reconciled.

Following data cleaning, all firm identifiers were removed, and variables were anonymized before release. CEO and HR director demographics (e.g., age, gender, tenure, education) were retained in coded form but without any personal identifiers. After cleaning, the dataset contained 508 valid cases with complete information across the four survey blocks and the integrated 2015 SABI financial indicators. Item-level missing values are coded as blanks in the dataset, and no imputation procedures were applied.

In order to correct minor deviations between the population and the achieved sample across the sector $\times$ age strata, post-stratification weights were computed once the dataset had been fully cleaned. The final dataset includes the variable *weight_ps*, which allows users to generate weighted estimates aligned with the sampling frame. All processing steps were implemented using standard statistical software. Post-fieldwork data processing, coding, and cleaning were conducted using Excel and SPSS, which were also used to prepare the final dataset, variable documentation, and supplementary files.

4.4. Descriptive statistics

The final dataset comprises 508 valid cases of family firms in industry and services drawn from the sampling frame constructed using the Family Business Firms Institute database. The sample is representative of the segment of identifiable medium-sized family firms included in this source, where all these firms are self-identified as family firms and meet the inclusion criteria. On average, firms employed 96 workers (median = 80) and reported operating revenues of €15.2 million (median = €8.5 million). Assets averaged €12.4 million, with wide dispersion (standard deviation = €18.6 million). These figures describe the distribution of firm size and financial resources in the sample.

Table 3
Descriptive statistics of family variables.

Variable	Mean	Median	SD	Min	Max
Family ownership (%) (*P2a_pro*)	95.61	100.0	12.46	35.0	100.0
Non-family ownership (%) (*P2b_pro*)	4.50	0.0	13.03	0.0	100.0
Family generation in charge (*P3_gen*)*	1.71	2.0	0.72	1.0	4.0
Total directors (*P4a_dire*)	3.87	3.0	3.30	1.0	50.0
Family directors (*P4b_dirf*)	2.20	2.0	1.40	0.0	8.0

Note: (*) Percentage of generations of family firms: 43% first, 44% second, 11% third, and 2% fourth or later.

4.4.1. Family dimension

As shown in Table 3, there is a strong concentration of ownership and family involvement in governance. On average, families hold 95.6% of equity, while non-family ownership is nearly negligible, which is consistent with the sample design and the central role of family control in Spanish SMEs. Regarding generational leadership, 43% of firms are led by the first generation, 44% by the second, 11% by the third, and 2% by the fourth generation or later. Boards of directors are small (mean = 3.9 members), with an average of 2.2 family members.

The SEW scale further illustrates the salience of family-centered, non-financial goals. The overall mean score (4.02 on a 1-5 scale) indicates that respondents generally attach high importance to socioemotional wealth. Items related to family unity, continuity, reputation, and stakeholder relationships are consistently rated as "important" or "very important." At the same time, variation across items reveals that not all firms prioritize SEW dimensions equally. This heterogeneity aligns with the view that family firms are not a homogeneous category but differ in the extent to which they emphasize control, identity, or social ties across generations. Such variation offers opportunities to explore how family-centered logics coexist with financial objectives, and how generational stage and governance structures condition the preservation of SEW.

4.4.2. HRM practices

The HRM block reveals substantial heterogeneity across firms. On average, companies trained 60 non-managerial employees in 2015 (median = 41), although the range extends from none to more than 1,600 employees. Training hours are equally dispersed, with a mean of 951 hours (median = 500; maximum = 16,000). Service-sector firms trained slightly more employees (mean = 66) than industrial firms (mean = 47), while third-generation firms tended to invest more in training than those led by other generations.

Responses to the 18-item Likert scale indicate a broad diffusion of HRM practices, which is consistent with the AMO framework. The most highly rated practices relate to ability (e.g., job-specific selection criteria, mean = 4.25), motivation (e.g., equity in compensation, mean = 4.16), and opportunity (e.g., participation in decision-making, mean = 3.6). Other practices, such as performance appraisal and communicated career plans, obtain mean scores between 3.7 and 3.3, suggesting moderate but widespread adoption.

Taken together, these results show that Spanish family firms are increasingly formalizing their HRM systems and adopting bundles of high-performance work practices. However, the variability observed across practices and firms also reflects the tension between traditional family-centered management and the implementation of structured, professionalized HRM systems.

4.4.3. Financial indicators

Survey responses are complemented by objective measures from SABI for the 2015 financial year. EBITDA averaged €1.1 million (median = €0.57 million), while equity averaged €6.3 million –again with strong dispersion. Profitability is moderate but heterogeneous: mean ROA is 4.5% (median = 3.6%), and mean ROE is 14.2% (median = 9.1%). Extreme values suggest that some firms face financial stress, whereas others achieve exceptional returns.

Taken together, these descriptive statistics portray a population of firms that are highly family-controlled, predominantly second-generation, and diverse in their HRM and financial systems. The heterogeneity documented in training policies, HRM practices, and profitability pro-

vides fertile ground for comparative analyses by sector, generation, or governance structures. Moreover, the alignment of subjective survey responses with objective financial indicators reinforces the dataset's potential for replication, validation, and meta-analytic research in the family business field.

5. Limitations

The dataset has some limitations that should be considered by users. A first limitation relates to non-response. Although the sampling followed a stratified random design, detailed call outcomes were not recorded, thus preventing a complete non-response analysis. To partially mitigate this issue, post-stratification weights were computed to correct sample-population discrepancies across strata. Although useful for improving representativeness, weighted analyses should be interpreted with caution, as weighting may increase sampling variance. Second, the data are cross-sectional, combining survey responses collected in 2016 with financial indicators from the 2015 fiscal year. Although this design ensures temporal alignment, some variables for activities may involve a significant recall window. Even though anchoring cues were used systematically and despite the fact that HRM practices tend to be relatively stable, users should remain aware of possible bias when interpreting these indicators. Third, although several procedural safeguards were implemented –such as standardized scripts, neutral wording, screening of respondents, and separation between questionnaire blocks– perceptual and social-desirability biases cannot be entirely ruled out. Users should be aware of this limitation –particularly with regard to subjective assessments of SEW and HRM practices. Fourth, the set of economic-financial indicators is restricted to basic measures such as employment, revenues, assets, EBITDA, equity, current assets and liabilities, ROA, and ROE. Researchers who require additional financial ratios (e.g., cash flows, debt maturity) will need to complement the dataset with further SABI extractions. Finally, the dataset focuses on medium-sized Spanish family firms in the industry and services sectors, which may limit generalizability to other family firm size categories or sectors. Despite these boundaries, the dataset does remain a unique empirical resource for family business research.

Ethics Statement

All respondents were fully informed about the objectives, scope, and voluntary nature of the study before participation. Verbal informed consent was obtained prior to the interview. The survey was carried out in accordance with the principles of the Declaration of Helsinki. No personal identifiers are included in the dataset, which is fully anonymized and released in compliance with data protection regulations. Given the non-interventional nature of the research (structured managerial survey and secondary financial data), formal approval by an institutional ethics committee was not required under Spanish regulations.

CRediT Author Statement

All authors contributed equally to this manuscript. Specifically, **J. Samuel Baixauli-Soler:** Funding acquisition, Supervision, Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. **María Belda-Ruiz:** Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. **Gabriel Lozano-Reina:** Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. **Juan David Peláez-León:** Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. **Gregorio Sánchez-Marín:** Project administration, Funding acquisition, Supervision, Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization.

Data Availability

Family firms: SEW, HRM practices, and financial indicators (Original data) (Zenodo).

Acknowledgements

This work was supported by Fundación Cajamurcia and by the Agencia Estatal de Investigación of the Government of Spain (project reference ECO2017-84209-P). In addition, it was supported by the project PID2024-159036NA-100, funded by MICIU/AEI/10.13039/501100011033 and by FEDER, EU.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] C. Arz, A. Kuckertz, Survey data on organizational culture and entrepreneurial orientation in German family firms, Data Br. 24 (2019) 103827, doi:10.1016/j.dib.2019.103827.
[2] A.L. Meroño-Cerdan, C. Lopez-Nicolas, F.J. Molina-Castillo, Drivers of business model innovation in family firms: a dataset analysis, Data Br. 54 (2024) 110443, doi:10.1016/j.dib.2024.110443.
[3] J.J. Chrisman, J.H. Chua, P. Sharma, Trends and directions in the development of a strategic management theory of the family firm, Entrep. Theory Pract. 29 (5) (2005) 555–576, doi:10.1111/j.1540-6520.2005.00098.x.
[4] E. Gedajlovic, M. Carney, J.J. Chrisman, F.W. Kellermanns, The adolescence of family firm research: taking stock and planning for the future, J. Manag. 38 (4) (2012) 1010–1037, doi:10.1177/0149206311429990.
[5] P. Berrone, C. Cruz, L.R. Gómez-Mejía, Socioemotional wealth in family firms: theoretical dimensions, assessment approaches, and agenda for future research, Fam. Bus. Rev. 25 (3) (2012) 258–279, doi:10.1177/0894486511435355.
[6] L.R. Gómez-Mejía, K.T. Haynes, M. Núñez-Nickel, K.J.L. Jacobson, J. Moyano-Fuentes, Socioemotional wealth and business risks in family-controlled firms: evidence from Spanish olive oil mills, Adm. Sci. Q. 52 (1) (2007) 106–137, doi:10.2189/asqu.52.1.106.
[7] B.J. Debicki, F.W. Kellermanns, J.J. Chrisman, A.W. Pearson, B.A. Spencer, Development of a socioemotional wealth importance (SEWi) scale for family firm research, J. Fam. Bus. Strategy 7 (1) (2016) 47–57, doi:10.1016/j.jfbs.2016.01.002.
[8] J.S. Baixauli-Soler, M. Belda-Ruiz, G. Sánchez-Marín, Socioemotional wealth and financial decisions on private family SMEs, J. Bus. Res. 123 (2021) 657–668, doi:10.1016/j.jbusres.2020.10.022.
[9] J.D. Peláez-León, G. Sánchez-Marín, Socioemotional wealth and human resource policies: effects on family firm performance, Int. J. Entrep. Behav. Res. 28 (1) (2022) 109–135, doi:10.1108/IJEBR-05-2021-0404.
[10] J.D. Peláez-León, G. Sánchez-Marín, High-performance work systems in family firms: a mixed gamble approach, J. Bus. Res. 156 (2023) 113532, doi:10.1016/j.jbusres.2022.113532.
[11] G. Sánchez-Marín, G. Lozano-Reina, M. Beglaryan, Linkages between high-performance work practices and family-centered goals: implications for financial performance in family firms, J. Small Bus. Enterp. Dev. 31 (1) (2024) 126–151, doi:10.1108/JSBED-03-2023-0117.
[12] K. Jiang, J. Hu, S. Liu, D.P. Lepak, Understanding employees' perceptions of human resource practices: effects of demographic dissimilarity to managers and coworkers, Hum. Resour. Manag. 56 (1) (2017) 69–91, doi:10.1002/hrm.21771.
[13] E. Appelbaum, T. Bailey, P. Berg, A.L. Kalleberg, Manufacturing Advantage: Why High-Performance Work Systems Pay Off, Cornell University Press, 2000.
[14] J.C. Casillas, M.C. López, A. Meroño-Cerdán, A. Pons, R. Baiges, La empresa familiar en España, Inst. Empresa Fam. (2015).
[15] D.I. Pielsticker, M.R. Hiebl, Survey response rates in family business research, Eur. Manag. Rev. 17 (1) (2020) 327–346, doi:10.1111/emre.12375.
[16] S.D. Kofford, W.G. Dyer, R.L. Sorenson, W.J. Dyer, J.M. Milbrandt, Family faith, altruistic values and community CSR in family firms, J. Fam. Bus. Manag. 15 (6) (2025) 1702–1720, doi:10.1108/JFBM-03-2025-0067.