



© Tanayoch-AdobeStock

LIGHTS - Future of HEI

Reviewing and classifying scientific contributions from ESCP Business School

ESCP Impact Paper No.2024-10-EN

Ivan SAVIN & Régis COEURDEROY
ESCP Business School

[Reviewing and classifying scientific contributions from ESCP Business School

Ivan Savin *
Régis Coeurderoy *

ESCP Business School

Abstract

This study explores main research themes addressed by the faculty of the ESCP business school. To this end, we utilize the structural topic modelling algorithm. 14 main topics ranging from energy markets and sustainability to entrepreneurship and consumer behaviour are identified. While topics related to leadership, technology and entrepreneurship were increasing in popularity over time, themes on financial markets and customer relationship were comparatively declining. We discuss geographical distribution of the revealed research themes and how the topics vary in terms of citations received and rank of journal they are published in.

Keywords: computational linguistics; literature review; machine learning; topic modelling.

*Professor, ESCP Business School

Reviewing and classifying scientific contributions from ESCP Business School

ESCP is the oldest business school in the world and one of the top Grande École de Management in France. ESCP has a unique structure with six urban campuses in major European cities - Paris, London, Berlin, Madrid, Turin and Warsaw – allowing it to teach and conduct research in several countries (Kaplan, 2014). Over the last couple of decades ESCP has considerably increased its research activity producing over 2000 peer-reviewed scientific publications. We believe it is a good time to take stock and see which themes have been a major focus by the faculty of the school and how these have historically evolved over time. The advantage of this approach is that we go beyond the sheer number of publications looking on the key areas ESCP has contributed scientifically.

Given the large number of articles published, reviewing them with conventional tools and human coders is very challenging. Fortunately, over the last years computational linguistics methods from the intersection of machine learning and natural language processing have been developed. Using these methods to review publications and analyse trends in certain outlets and research fields has become popular in the literature (Griffith & Steyvers, 2004; Mo et al., 2015; De Battisti et al., 2015; Lüdering & Winker, 2016; Ambrosino, et al., 2018; Maier et al., 2018; Asmussen & Møller, 2019; Savin 2023, Borah et al., 2023). We apply the topic modelling technique to titles, abstracts and keywords of articles indexed in the Scopus database until October 2023. More specifically, we derive 14 main themes¹ by applying structural topic modelling method. Subsequently, we focus on temporal and geographical dimensions of the identified 14 topics, and which of them have attracted more citations or have been published in top-ranked journals in management and economics. We exclude editorials and errata, as well as studies published not in the top tiers of the ESCP Journal List ending up with a total of 792 research articles.

Data and Methods

The sample of studies published by researchers with ESCP affiliation has been retrieved from the Scopus database² on October 8, 2023, and included 2027 studies. For this exploratory study, the choice was made to start from the ESCP Journal List (edition 14th of February 2023). The list may have an influence on the choices by ESCP researchers as it is articulated with a bonus policy. The Journal List is composed of external official lists from international and European national lists of journals in management and economics, namely Financial Times, FNEGE, ABS, VHB, and CNRS. See Rodrigo et al. (2024) for more details on the correspondence of the List with other official journal rankings. Naturally, this list primarily focuses on journals which are of major importance for business schools around the world, ESCP included. The list covers over 2000 journals ranked from D (the lowest rank) through C, B, A and AFT³ to ALPHA (the highest position). Over 600 journals in the ESCP list have a rank B and higher.

¹ Henceforth, we use the words “theme” and “topic” as equivalents.

² Scopus database launched by Elsevier, next to Web of Science from Thompson Reuters, is among the most widely used index and citation databases. We opted to use Scopus since it covers a considerably larger number of journals (Mongeon and Paul-Hus, 2016) and is easier to navigate (Burnham, 2006; Pranchkute, 2021).

³ AFT is a rank assigned to 9 out of 50 journals selected by Financial Times which have lower national ranks from lists like CNRS and VHB, which places them lower than ALPHA (remaining 41 Financial Times journals). An example of ALPHA journal is Academy of Management Journal, while an example of AFT is Harvard Business Review.

We filtered the sample in the following ways. First, we removed studies published in layouts not listed in the ESCP journal list (852) and publications having appeared in journals ranked as C (268) or D (76). This has been done to ensure that we focus only on the contributions of highest quality and devoted to the topics of economics and management.⁴ Second, we excluded from further consideration 40 editorials and errata. The reason for this is twofold: first, those studies typically contain no abstract⁵ and thus would be very hard to analyse using our method of textual analysis. Second, these papers typically provide an overview of or correct already existing publications without contributing new knowledge. As a result, we obtained the distribution of publications over time and by rank in the ESCP journal list demonstrated in Figure 1. As one can see, the number of publications was steadily rising over time, reaching 98 in 2023. The figures show the strategic inflexion of ESCP towards research in the last decade. More than half of the articles (530) are published in B journals, 151 in A outlets, and 111 in ALPHA or AFT layouts. Also Figure 1 shows that the proportion of higher quality journals (ALPHA and AFT) has increased in the last decade and approaches 20%.

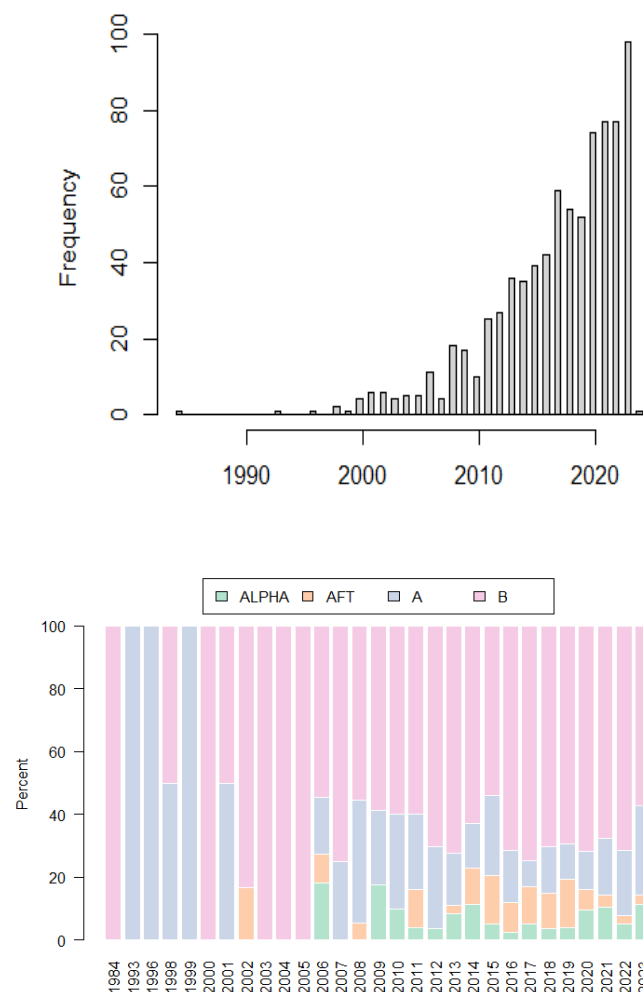


Figure 1. Number of studies published over time in our sample and their distribution by rank in the ESCP list.

⁴ While some high-quality like Nature are not in the ESCP list, those journals are focusing on interdisciplinary, environmental and other topics, which are not the core focus of ESCP as a business school. We have experimented with the whole sample of studies as well using topic modelling. Results are available on request.

⁵ For some publications in *Revue Française de Gestion* no abstract was available in the Scopus database, so we manually retrieved them from the website of the journal in French language and translated using Google Translate.

Table 1 lists three journals that appear most often in the sample for each of the four ranks in the ESCP Journal List. Since B journals comprise over 50% of our sample, it is not surprising that the most popular outlets in the whole sample are from this rank: Journal of Business Research and European Management Journal. Articles published by ESCP faculty in Journal of Business Research and in Journal of Marketing are cited considerably more often than publications in other popular outlets (see last column in Table 1). Revue Francaise de Gestion is the only non-English journal in Table 1 and has much fewer citations, which can be explained by a smaller audience of the French-speaking journal.

Table 1. Journals with highest number of publications in each rank in our sample

Journal	Rank	Number of studies published	Average number of citations per year
Organization Studies	ALPHA	7	3.38
Journal of Marketing		5	14.23
Journal of Business Venturing		3	5.49
Journal of Business Ethics	AFT	31	6.36
Human Resource Management		13	7.15
Human Relations		4	4.91
International Journal of Research in Marketing	A	14	12.98
European Journal of Operations Research		13	5.09
International Journal of Operations and Production Management		10	4.37
Journal of Business Research	B	56	14.23
European Management Journal		49	2.99
Revue Francaise de Gestion		32	0.59

If we look on the coverage of countries where scientists from our sample were affiliated⁶ (note that here we look on all co-authors of the 792 articles and not only on those affiliated at ESCP), we will find a list of 77 countries (top plot in Figure 2) with the five countries where ESCP has campuses⁷ and the US being the top states, and France leading by a big margin, which is not surprising given that the campus in France is by far the largest in the school. If we weight countries contribution by the number of citations per year the respective studies have received (bottom chart in Figure 2), the colour palette changes marginally: the most notable difference is that Germany and the US increase their weight. The Figure 2 thus provides a good overview on the geographical coverage of research collaborations of ESCP faculty showing a global research profile. Later in the manuscript we will return to the issue of country affiliation when we analyse the topics.

To discover latent structure of our textual data, we use topic modelling. Putting it simply, TM clusters words into topics based on how often any pair of words appears in the same texts (Blei, 2012; Savin and van den Bergh, 2021). For example, words “innovation”, “technology”, “knowledge”, and “open” standing next to each other in one of the topics we report in the next section implies that these words appear considerably more often in combination with each other and other words from this topic. This way TM looks on words not in isolation, but accounting for their context, which can influence their meaning. An advantage of structural topic modelling (STM) over classical TM is that it includes additional

⁶ Many articles in our sample have authors affiliated in more than one country. To quantify share of a country, we identify the full list of unique countries mentioned in affiliations of each publication and then attribute equal shares to each of the countries. For example, if a study had two co-authors with one affiliated to Spain and the other to Spain and Germany, each country would receive 0.5 weight in our counting. The reason for such a simplistic rule is that Scopus does not provide affiliations for each author separately, but a list of affiliations put together.

⁷ The ESCP campus in Warsaw has been opened relatively recently and compared to other five urban campuses it is the smallest one, which explains so far moderate contribution of Poland in the scientific output of ESCP.

data about the publications (like year of publication and average number of citations per year received), which increases quality⁸ of the topics produced (Roberts et al., 2014).

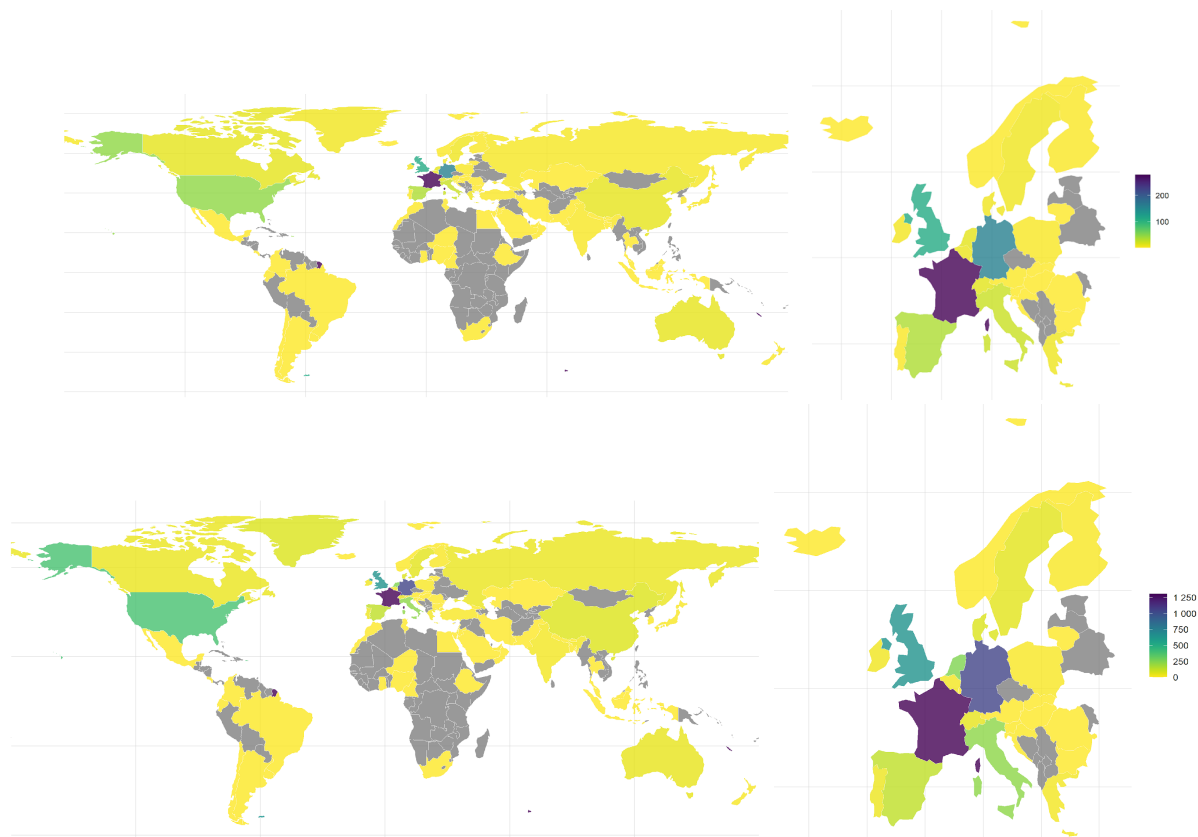


Figure 2. Country coverage based on affiliations stated in our sample.

Note: The top heatmap shows the global coverage where the studies with authors affiliated at more than one unique country are equally split between those countries. The bottom heatmap additionally takes the number of citations received per year into account.

Before applying STM, we did some standard steps described in the literature to clean and prepare the data (Aggarwal, 2018; Uglanova & Gius, 2020; Savin and Teplyakov, 2022). Specifically, we replaced capital letters; removed punctuation and stop words; transformed words to their dictionary form using lemmatization; removed words that appear in less than three publications; and identified stable expressions as single words (bigrams), e.g., “supply_chain”, “decision_making” using the normalized pointwise mutual information (NMPI) score (Bouma, 2009). Accordingly, our final dataset contains 2462 unique words and 51614 total word occurrences for building a topic model.

To determine appropriate topic number, we run STM for 3 to 50 topics and record model performance on the following common metrics (Roberts et al., 2019): heldout log-likelihood (i.e. predictive accuracy), exclusivity (to what extent most popular words in the topics are not overlapping), and semantic coherence (co-occurrence of words from the same topic in text documents). Naturally, larger topic number increases the model's predictive power and topic exclusivity but reduces their semantic coherence. In Figure A1 in the Appendix, we demonstrate that 14 topics allow us to achieve maximum possible predictive accuracy and close to the maximum exclusivity while maintaining semantic coherence at reasonable level.

⁸ Roberts et al. (2014) define major qualities of topics as follows: they are cohesive in the sense that high-probability words for the topic tend to co-occur within documents, and they are exclusive in the sense that the top words for that topic are unlikely to appear within top words of other topics.

Results

The resulting themes are presented in Table 2 and visualised as word clouds in Figure 3. Furthermore, Figures 4 and 5 summarize topic dynamics over time and frequency of citations. In Table 2 next to most frequent and exclusive words for each topic and an illustrative title of a paper with the highest prevalence of the topic we provide concise topic labels we have formulated after studying titles, abstracts and keywords of top twenty documents with highest prevalence of the respective topics.

Topic 1 (henceforth, Tx stands for topic x) on theoretical framework has the largest prevalence in our sample of documents (11.7%) and is most generic covering conceptual papers and literature reviews, among others. This topic was relatively stable over time (Figure 4), while topics 2, 3 and 10 on leadership, technology and entrepreneurship have significantly increased in their popularity, while topics 8 and 13 on financial markets and customer relationship, to the contrary, have relatively decreased⁹. Other topics, just like T1 have been relatively stable over time.

If we look on the propensity to get cited (Figure 4), the top topics are T3 on technology, T5 on sustainability and T13 on customer relationship considerably exceeding the average number of citations a paper in our sample was receiving (5.9 citations per paper per year). Among topics which lag behind on this indicator are T11 on fiscal and monetary policy and T4 on social responsibility of business.

If we look on the distribution of topic by quality of the journal they were published in based on the ESCP list (Figure 6), we will find that T3 on technology and innovation is more common among ALPHA journals, T9 and T11 on information systems and fiscal and monetary policy – among AFT journals, while T12 and T14 on energy markets and branding and culture – among A journals. With a range between 11.7% and 5.3% and a C4 ratio¹⁰ at 38.8%, the profile of ESCP is rather diversified, as we can expect from a large generalist business school.

⁹ Note that in Figure 4 we are looking on the share of each topic per year. If we instead, plot the cumulative number of studies published on each topic by the ESCP faculty (see Figure A2 in the Appendix), we will find that the scientific contribution by ESCP has increased in virtually all areas.

¹⁰ C4 ratio is a common concentration ratio measuring the total share of the four largest items. The higher the ratio, the more concentrated the research agenda of the ESCP faculty.

Table 2. 14 main topics based on titles, abstracts, and keywords.

	Topic label	Most discriminating terms and illustrative titles	Topic proportion
T1	Theoretical framework	review, field, academic, agenda, systematic, overview, scholar, conceptual, framework, synthesis "How to advance theory through literature reviews in logistics and supply chain management"	11.7%
T2	Leadership and work practices	employee, human_resource, career, leadership, talent, top, job, diversity, compensation, team "The effect of HRM attributions on emotional exhaustion and the mediating roles of job involvement and work overload"	11.0%
T3	Technology and innovation	innovation, capability, technological, orientation, technology, digital_transformation, transformation, knowledge, open, industry "Joining forces or going it alone? On the interplay among external collaboration partner types, interfirm governance modes, and internal R&D"	8.1%
T4	Social responsibility of business	csr, institutional, network, informal, actor, accountability, society, social, intervention, institution "Neoliberalism, crises and accusations of fraud: A vicious circle of reinforcing influences?"	8.0%
T5	Sustainability	sustainability, stakeholder, circular_economy, ethical, project, sustainable, creation, resource, value, reporting "The Organizational Dynamics of Business Models for Sustainability: Discursive and Cognitive Pathways for Change "	7.4%
T6	Consumer behaviour	decision_making, consumption, evaluation, behavior, user, willingness_pay, consumer, choice, purchase "Never too rich to care about prices: Effects of price endings on customer perceptions of luxury"	6.9%
T7	Corporate value	shareholder, foreign, corporate, ownership, governance, investor, stock, trade, reputation, dividend "Common ownership and firm dividend policies"	6.9%
T8	Financial markets	commodity, volatility, option, price, crude_oil, variance, index, portfolio, diversification, inflation "General multilevel Monte Carlo methods for pricing discretely monitored Asian options"	6.3%
T9	Information systems	inventory, rfid, item, identification, iot, control, real_time, deterioration, information, ticket_switch "RFID-enabled flexible warehousing"	6.1%
T10	Entrepreneurship	entrepreneurship, entrepreneurial, venture, entrepreneur, effectuation, motivation, attitude, intention, commitment, nascent "Teaching proactivity in the entrepreneurial classroom"	5.8%
T11	Fiscal and monetary policy	bank, fiscal, credit, banking, tax, debt, monetary_policy, shock, government, regulation "Income inequality and monetary policy in the euro area"	5.6%
T12	Energy markets	public, energy, private, electricity, environmental, efficiency, country, performance, profitability, improvement "Measuring the efficiency of energy-intensive industries across European countries"	5.5%
T13	Customer relationship	supplier, customer, marketing, advertising, buyer, resilience, relationship, disruption, supply_chain, crm "Capturing value creation in business relationships: A customer perspective"	5.4%
T14	Branding and culture	brand, social_medium, identity, cross_cultural, culture, attachment, content, luxury, digital, fashion "The effect of culture on perception and cognition: A conceptual framework"	5.3%



Figure 3. Word clouds encompassing 100 most frequent and exclusive words for each of the 14 topics in ESCP research.

Note: The font size reflects the probability of the respective word given the topic, while darker colour indicates higher exclusivity.

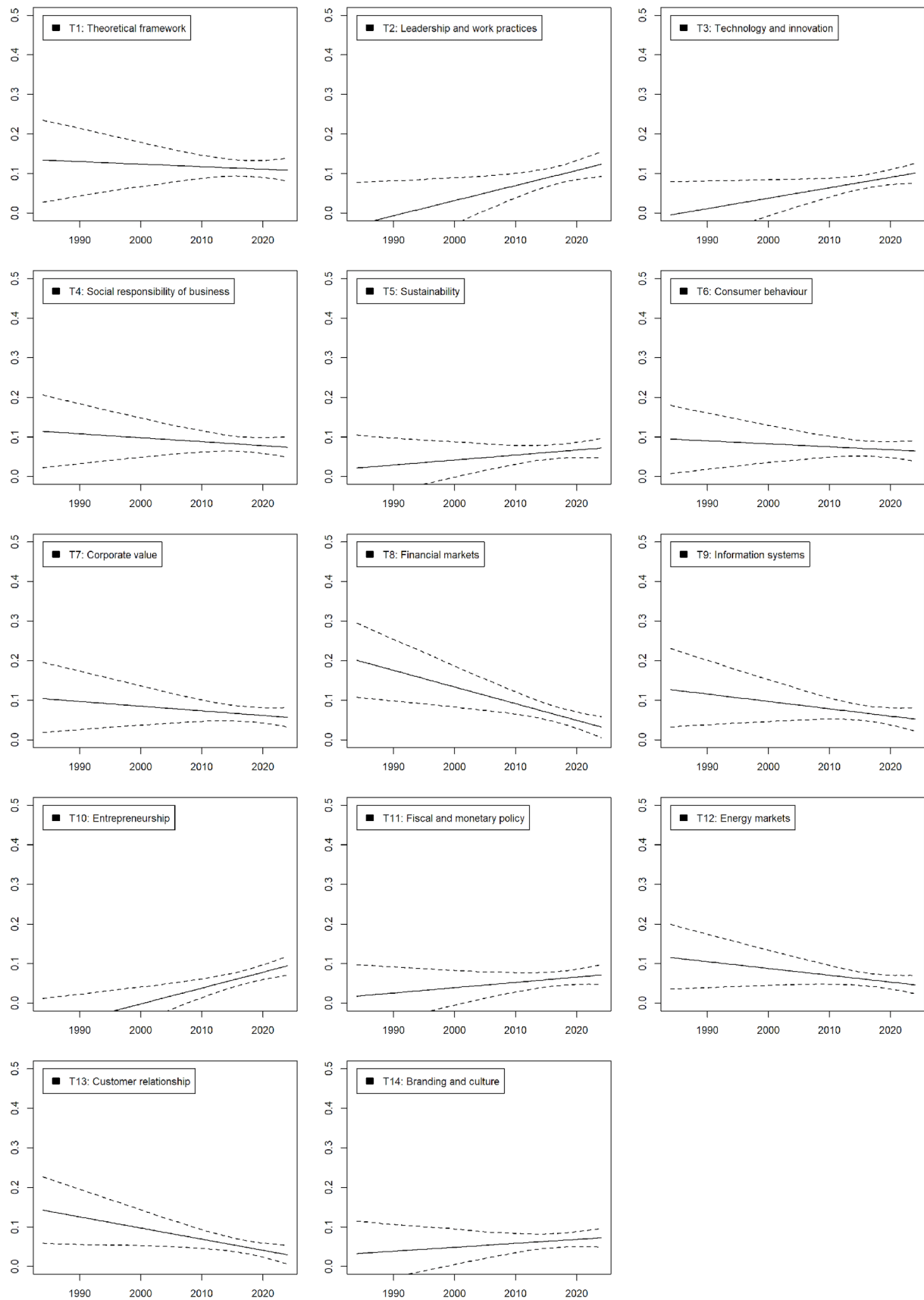


Figure 4. Change in shares of topics in ESCP research over time.

Note: To establish presence of a relation between the period of publication and the topic prevalence, we fit a linear regression model, where the outcome variable is the topic prevalence in each publication. An estimate with a 95% confidence interval above zero indicates a significant prevalence of that topic in the respective year.

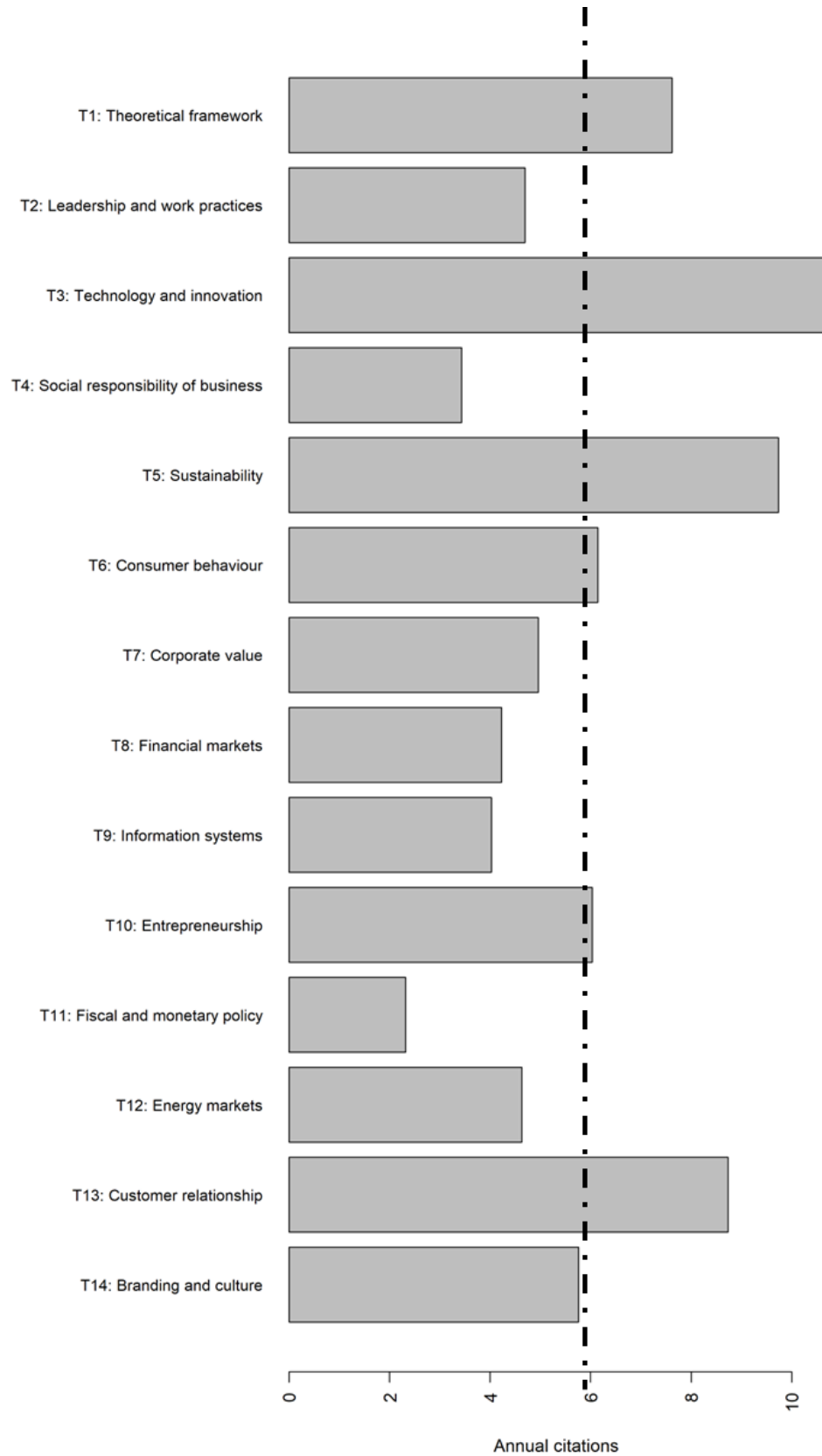


Figure 5. Average number of citations per year per document for each topic.

Note: The number of citations per document belonging to a topic is calculated by multiplying the publication's number of citations per year with the topic prevalence and taking the average. The dashed line shows the average number of citations per year in our sample (5.93).

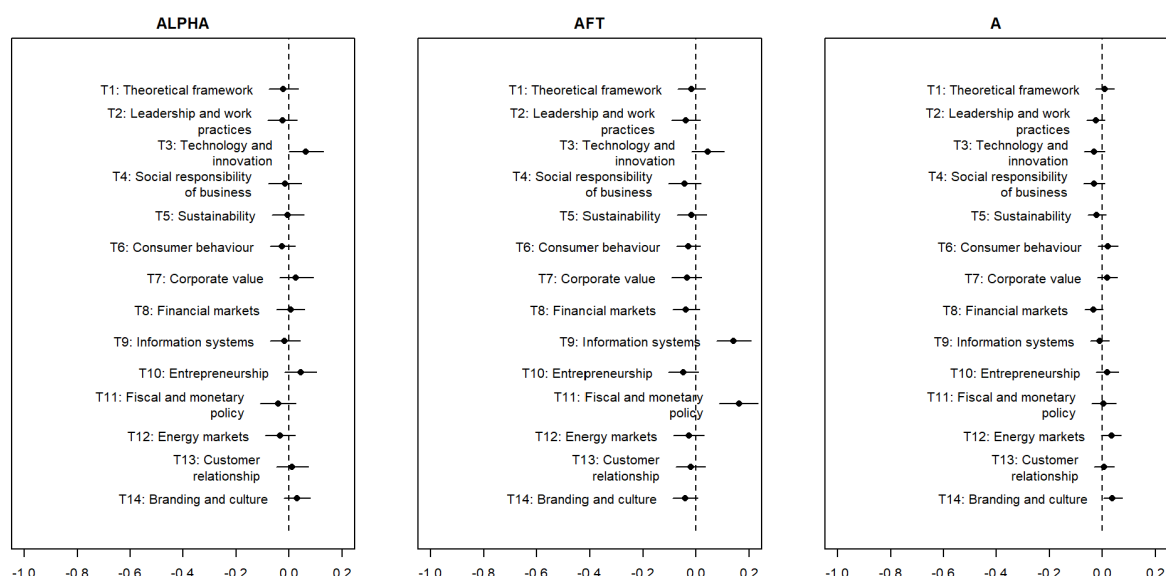


Figure 6. Distribution of topics by rank in ESCP journal list.

Note: The plots show mean difference in topic proportions between the rank under consideration and rank B which is used as a benchmark (a positive value on the X-axis indicates a larger prevalence in the corresponding region). In other words, Figure 6 is a result of a linear regression where topic prevalence is the dependent variable and ranks higher than B are independent (explanatory) dummy variables.

If we look at which topics researcher affiliated in different countries tend to focus (Table 3), we will notice that with authors from France contributing most to topics 1 and 4 (theoretical framework and social responsibility). The London campus and the UK more broadly also focus on T1 but also T10 (financial markets). Authors affiliated in Germany have a strong specialization on T2 about leadership and work practices. The campuses in Spain and Italy have some similarity in their core attention to topics 3 (technology and innovation) and 10 (entrepreneurship). Thus, we observe some specialization effects in campuses with a more limited size.

Table 3. Distribution of the main topics across six main countries.

	France	Germany	UK	Spain	Italy	US
T1 Theoretical framework	14	9.9	10.4	6.7	7.3	8.8
T2 Leadership and work practices	7.4	20.3	10	6.6	7.7	5.1
T3 Technology and innovation	6.2	9.1	8.1	15	30.8	5
T4 Social responsibility of business	11.8	4.8	7.3	3.5	3.4	6.5
T5 Sustainability	7.3	8.7	7.2	4.7	11.2	5.1
T6 Consumer behaviour	5	8.5	8.8	10.8	1.6	6.2
T7 Corporate value	7	7.8	7.6	4.2	4.2	6
T8 Financial markets	8.2	2.8	10.2	4.6	5.7	3
T9 Information systems	7.1	3.2	1.8	1.8	0.7	23.6
T10 Entrepreneurship	4.6	6.9	5.2	12.8	11.6	5.4
T11 Fiscal and monetary policy	7.8	2.5	5.5	8.1	3.2	6.3
T12 Energy markets	5.3	4.3	7.2	5.3	3.2	5.3
T13 Customer relationship	4.5	8.1	2.6	6.5	6.2	7
T14 Branding and culture	3.9	3.2	8.1	9.5	3.2	6.7

Note: Results are provided in percentages so that each column (country) sums up to 100. Colour corresponds to the percentage value (ranging from red for large values to blue for small ones).

Note that if we normalize the values in Table 3 not by the research output of each campus, but by the output on each topic, we find that on all the major topics the Paris campus contributes between 30-60% of the publications being leader in all topics except T2 and T13, while the campus in Turin on all topics except topic 3 contributes less than 10% of the publications (see Table A1 in the Appendix).

Conclusion

To conclude, we identified 14 main themes published by scholars with ESCP affiliation over the last few decades ranging from such broad topics like T1 on theoretical framework to more specific like T12 on energy markets and T14 on branding and culture. We found that attention to topics related to leadership, technology and entrepreneurship was rising over time, while topics on financial markets and customer relationship were comparatively declining. In terms of quality of publications (citations, rank of the journal by ESCP), a clear leader so far seems to be topic 3 on technology and innovation. This topic attracted on average largest number of citations and tends to be published more often in ALPHA journals. This trend has certainly been a key one in the last two decades with a profound transformation of businesses driven by the digital revolution. Finally, we discuss geographical distribution of the revealed research themes demonstrating that different ESCP campuses tend to focus on different issues thus complementing each other and contributing to the overall excellence of the business school. This exploratory work gives an image of a generalist higher education institution in the field of management at first. It also shows that the ESCP researchers deal with big issues for the world of business while maintaining a strong emphasis on theoretical foundations.

It is worth stressing that the results discussed in this paper are based solely on open-access information about publications retrieved from the Scopus database and the ESCP Journal List which is available online. The paper demonstrates the usefulness of tools based on machine learning in reviewing and classifying large amount of data in a convenient and easily accessible way thus advancing quality and transparency of scientometric analysis.

This exploratory study has some limitations. First, it is based upon the latest ESCP list, and the list has been modified several times over the last few decades. Secondly, we are aware that the choice of ALPHA, AFT, A and B journals is a truncated vision of the global research activities at the ESCP business school. Third, our study provides complementary, information and insights to traditional – and inevitably more subjective – qualitative review methods based on human coding. It does not substitute qualitative analysis as it was not aimed to answer deeper questions such as on main findings of the articles, even though recent advances with large language models such as ChatGPT demonstrated some potential of computer-based algorithms for more nuanced analysis. We leave this kind of questions for future research.

Our results can serve as guidance for future contributors of ESCP faculty helping to derive ideas on which topics attracted more attention in the past and which only recently became popular, and colleagues from which country/campus tend to specialize on similar ideas.

Appendix

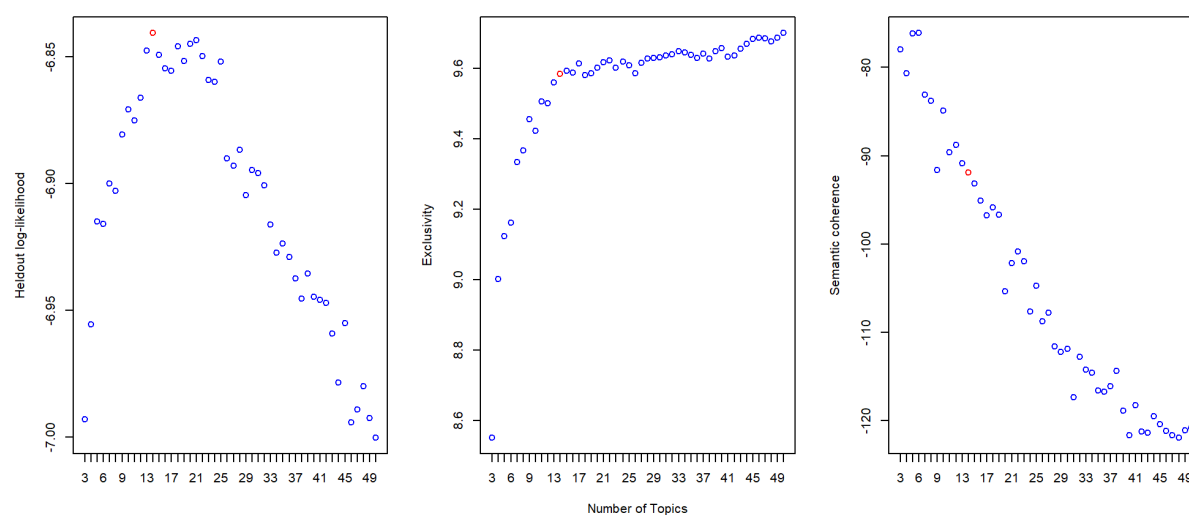


Table A1. Distribution of the main topics across six main countries.

	France	Germany	UK	Spain	Italy	US
T1 Theoretical framework	52.5	21.3	15.3	3.1	2.2	5.5
T2 Leadership and work practices	29.2	46.1	15.6	3.2	2.5	3.3
T3 Technology and innovation	31.3	26.3	16.1	9.3	12.7	4.2
T4 Social responsibility of business	61.4	14.2	15	2.3	1.5	5.6
T5 Sustainability	42	28.4	16.1	3.3	5.3	4.9
T6 Consumer behaviour	31.6	30.8	21.8	8.4	0.8	6.5
T7 Corporate value	42.9	27.4	18.3	3.2	2.1	6.1
T8 Financial markets	53.6	10.5	26	3.7	3.1	3.2
T9 Information systems	51.5	13.1	5	1.6	0.4	28.4
T10 Entrepreneurship	32.7	28.3	14.6	11.2	6.8	6.4
T11 Fiscal and monetary policy	56.9	10.5	15.8	7.3	1.9	7.7
T12 Energy markets	42.9	19.9	22.7	5.2	2.1	7.2
T13 Customer relationship	35.4	36.7	8.2	6.4	4	9.3
T14 Branding and culture	33.9	15.8	27.9	10.3	2.3	9.7

Note: Results are provided in percentages so that each row (topic) sums up to 100. Colour corresponds to the percentage value (ranging from red for large values to blue for small ones).

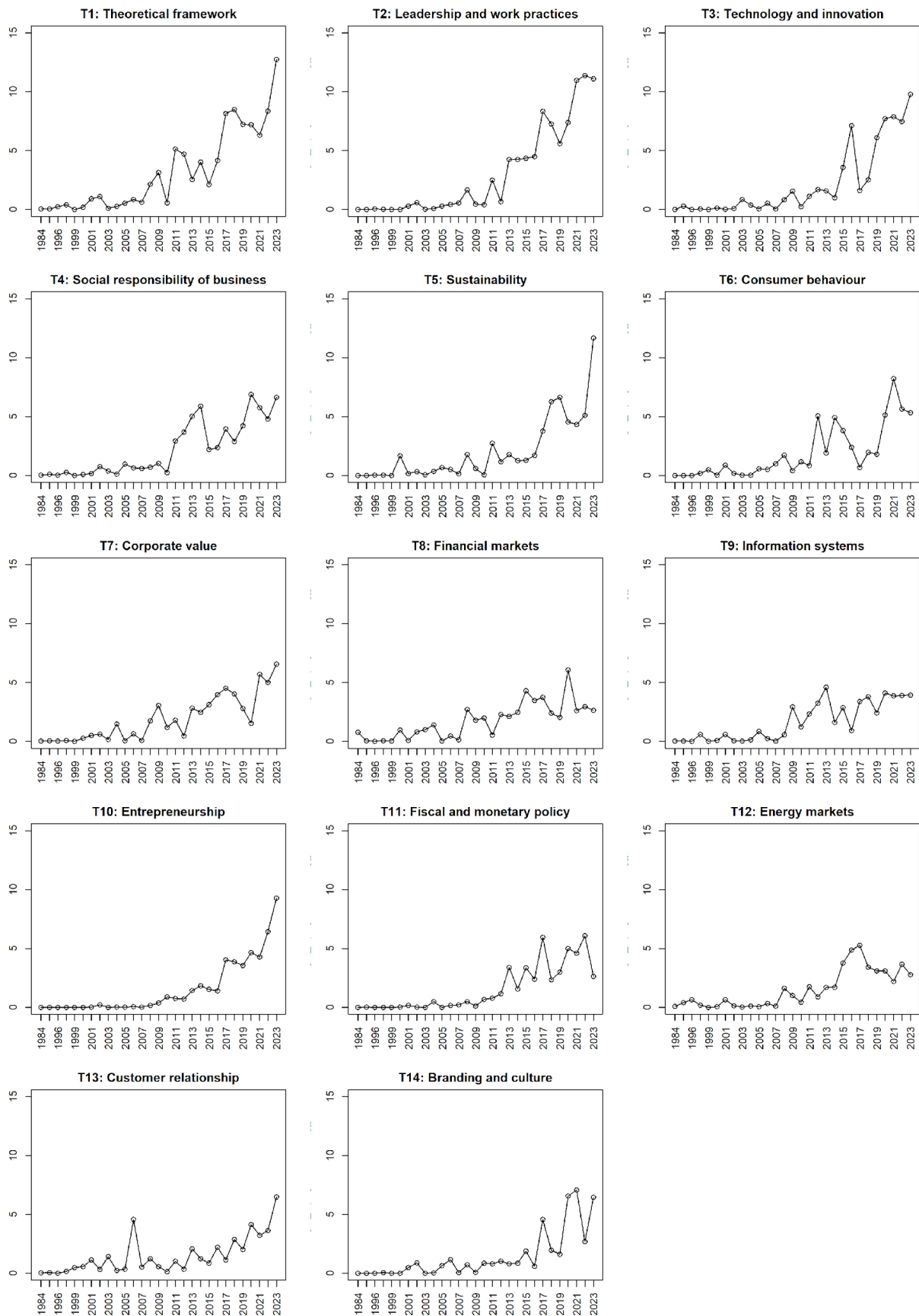


Figure A2. Number of publications per year and per topic.

Note: Since STM probabilistically assigns each paper typically to more than one topic with weights between 0 and 1 (topic prevalences), in this figure we cumulated the topic prevalences over time. That means that in reality there are more than 15 papers published by ESCP faculty on T1 in 2023, but since these papers encompass more than just this topic, cumulating the prevalences corresponding to T1 results in a value less than 15,

Références

- Aggarwal, C. C. (2018). Text Preparation and Similarity Computation. In: Machine Learning for Text (pp. 17–30). New York: Springer.
- Ambrosino, A., Cedrini, M., Davis, J., Fioria, S., Guerzoni, M., & Nuccio, M. (2018). What topic modelling could reveal about the evolution of economics. *Journal of Economic Methodology*, 25(4), 367–377
- Asmussen, C. B., & Møller, C. (2019). Smart literature review: a practical topic modelling approach to exploratory literature review. *Journal of Big Data*, 6(1), 1–18.
- Blei, D. M. (2012). Probabilistic Topic Models. *Communications of the ACM*, 55(4), 77–84.
- Borah, A., Bonetti, F., Calma, A., & Martí-Parreño, J. (2023). The Journal of the Academy of Marketing Science at 50: A historical analysis. *Journal of the Academy of Marketing Science*, 51(1), 222–243.
- Bouma G. (2009). Normalized (pointwise) mutual information in collocation extraction, *Proceedings of the International Conference of the German Society for Computational Linguistics and Language Technology*, GCSL 09.
- Burnham, J. F. (2006). Scopus database: a review. *Biomedical Digital Libraries*, 3(1), 1–8.
- De Battisti, F., Ferrara, A., & Salini, S. (2015). A decade of research in statistics: a topic model approach. *Scientometrics*, 103(2), 413–433.
- Griffith, T., & Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National Academy of Sciences of the United States of America*, 101, 5228–5235.
- Kaplan, A. (2014) European management and European business schools: Insights from the history of business schools. *European Management Journal* 32(4): 529–534.
- Lüdering, J., & Winker, P. (2016). Forward or backward looking? The economic discourse and the observed reality. *Jahrbuecher fuer Nationaloekonomie und Statistik* 236(4), 483–515.
- S. (2018). Applying LDA Topic Modeling in Communication Research: Toward a Valid and Reliable Methodology. *Communication Methods and Measures*, 12(2-3), 93–118.
- Mo Y., Kontonatsios, G., & Ananiadou, S. (2015). Supporting systematic reviews using LDA based document representations, *Systematic Reviews*, 4(1).
- Mongeon, P., & Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: a comparative analysis. *Scientometrics*, 106, 213–228.
- Pranckutė, R. (2021). Web of Science (WoS) and Scopus: The titans of bibliographic information in today's academic world. *Publications* 9(1), 12.
- Roberts, M. E., Stewart, B. M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S. K., Albertson, B., & Rand, D. G. (2014). Structural topic models for open-ended survey responses. *American Journal of Political Science*, 58(4), 1064–1082.
- Roberts, M. E., Stewart, B. M., & Tingley, D. (2019). STM: An R Package for Structural Topic Models. *Journal of Statistical Software*, 91(2), 1–40. DOI: 10.18637/jss.v091.i02.

Rodrigo, L., Ortiz-Marcos, I., & Palacios, M. (2024). A typology of social innovation: A comparative study of clustering methodologies. *International Entrepreneurship and Management Journal*, forthcoming

Savin, I., (2023). Evolution and recombination of topics in Technological Forecasting and Social Change. *Technological Forecasting and Social Change*, 194, 122723.

Savin, I., & Teplyakov, N. (2022). Using computational linguistics to analyse main research directions in Economy of Regions, *Economy of Regions* 18(2): 338-352.

Savin, I., & van den Bergh, J. (2021). Main topics in EIST during its first decade: A computational linguistic analysis. *Environmental Innovation and Societal Transition*, 41:10–17.

Uglanova, I., & Gius, E. (2020). The Order of Things. A Study on Topic Modelling of Literary Texts. In: *CHR 2020: Workshop on Computational Humanities Research* (pp. 57–76). Amsterdam, The Netherlands.