A Quantile-Based Composite Indicators Approach on Woman's Entrepreneurship as measured by Google Search Activity

C. Drago

Abstract Women's entrepreneurship research has exploded in popularity in recent decades. Therefore, women's entrepreneurship is considered a critical component of economic growth. Women's entrepreneurship faces many barriers. However, research shows that women's entrepreneurship is successful. Given this, we predict the future intentions of women entrepreneurs. Online search activity also provides relevant information about entrepreneurial intentions. Examining the entrepreneur's interactions with the government sheds light on how the government has supported her. So we investigated search patterns for female entrepreneurs through relevant Google queries for 2021 at the regional level for Italy. In this sense, we compared approaches and strategies by analyzing search trends in each region. Using Google searches, we produced a composite indicator based on quantiles and compared it with an interval-based composite indicator. The outcome is an indicator of women's interest in entrepreneurial activity. According to our analysis, interest in women's entrepreneurship is greater in four Italian regions, possibly due to better dissemination of helpful information. Interest in women's entrepreneurship has increased, and Google search activity is a predictor or "thermometer" of female entrepreneurship. The relevance of this work is also comprehensive of the possibility of exploring how the interest changes, varying the informational environment in which the woman entrepreneur works. At the same time, we can also generalize these results to understand women's entrepreneurial activities in different regional contexts. For example, Google searches may show greater engagement and interest in potential government help initiatives at the national or regional level. Finally, analyzing and assessing a woman's entrepreneurial activities based on search activity can be a valuable predictor of entrepreneurship. Identifying four well-performing regions means observing the decision-making process that leads to practical entrepreneurial decisions. In addition, the research can identify challenges, barriers, and solutions for women entrepreneurs to overcome and build their entrepreneurial projects. In this

C. Drago (✉)
Department of UNISU, Niccolò Cusano University, Rome, Italy
e-mail: carlo.drago@unicusano.it

389

respect, examining the interaction between women entrepreneurs and institutions is crucial.

1 Introduction

According to statistics,[1] the number of women starting a business has increased dramatically in recent decades (Wang 2013; Mayer 2008; Carranza et al. 2018). In addition, there is a growing literature on women entrepreneurs (Paoloni and Demartini 2016; Minniti 2010; Cardella 2020). Several studies prove this tendency and help women better understand the difficulties of becoming self-employed. Therefore, we can conclude that women's entrepreneurship is crucial for promoting economic growth in developing countries (Minniti 2010; Sharma 2018) and in general. Unfortunately, many obstacles, such as the business climate, stand in successful women's entrepreneurship (Dal Mas et al. 2019; Bianco et al. 2017). However, many problems and challenges can be overcome, such as the lack of know-how and business networks (Moletta et al. 2021). In this sense, relational and human capital is fundamental to successfully conducting entrepreneurial projects (Paoloni and Dumay 2015; Demartini and Paoloni 2011; Paoloni and Demartini 2012). Furthermore, in a growing economy, women need equal opportunities to start businesses since many businesses rely on outside capital to survive. Therefore, women should be allowed to succeed in a growing economy.

In this context, the intention and the interest in starting a business are fundamental. So the paper aims to measure women's entrepreneurial intentions in Italy, comparing the different regions. The comparison is performed by tracking Google Trends data to measure the relevant interest considering related queries useful to predict and forecast the economic phenomenon we are trying to measure (see Nagao et al. 2019; Choi and Varian 2012).

In several different ways, this study contributes to the literature. First, we explore composite indicators to measure women's interest in entrepreneurial activities. This indicator can have significant business implications because it can show relevant barriers to entrepreneurship and how they can be overcome. Secondarily the analysis can also show relevant experiences which can help the women's entrepreneurial activity. Thirdly, an approach based on quantile composite indicators is relevant because it allows more robust measurements related to classical composite indicators. For the rest of this article, we will follow the structure below. The following section reviews and discusses the literature on Google Trends data and latent phenomenon prediction. In the following part, we describe the methodology of

---

quantile-based composite indicators. Then, we discuss the results and identify some relevant policy implications. Finally, we review and discuss the approach's weaknesses and future developments. In order to measure the entrepreneur's intentions and interests, we have used Google Trends, which allows us to predict social phenomena (Cockcroft and Russell 2018).

Since its introduction, Google Trends has provided a new way of looking at search data that may show societal changes (Matias 2013). Furthermore, new research papers have shown the use of Google Trends to predict upcoming events successfully (Choi and Varian 2012; Matias 2013). Google Trends provides information on various topics and is a nowcasting tool. In this respect, following Donadelli and Gerotto (2019), it is possible to observe that specifically can exist a relationship between the frequency of Google searches and economic activity. In this respect, it is possible to observe also in Choi and Varian (2012) that there is an implied impact of the searches on the economic phenomena. So, it is possible to forecast different economic phenomena (the authors, for instance, predict unemployment claims). This work uses the different searches performed to evaluate the interest in entrepreneurial economic activities and predict women's entrepreneurial intentions in this context.

2 Methodology

In order to measure the entrepreneurial economic interest, we consider different searches on Google. Google Trend proved to be very useful for tracking economic activity (Donadelli and Gerotto 2019; Gómez Martínez 2015), so we concentrate on the different Google searches related to women's entrepreneurship. More specifically, the different queries are chosen to evaluate the relevant ones that can be searched by a woman entrepreneur or relevant in this context.

Furthermore, the searches cover widely the phenomenon we try to measure (interest in woman entrepreneurship in Italy). Of course, the choice of searches is an essential part of the analysis, so the highest number of searches is considered for constructing the composite indicator. The queries performed are enumerated in Appendix. The search period is from 2021-01-01 to 2021-12-31, so we covered the year 2021. The data considered covered Italy and considered queries in the Italian language. The data source in this sense is Google Trends (Google Trends 2022).

In order to combine the information related to the different searches over time, we have considered a composite indicator allowing for the use of all the different measurements using different queries. For example, the composite indicator measures the extent of interest in women's entrepreneurship and the related regulations or financing aspects. A composite indicator allows combining the information related to different variables, which are the different queries (related to different aspects of the phenomenon we are investigating). Composite indicators are essential statistical tools that can be used to measure complex phenomena. They also allow us

to construct rankings that can be used for policy purposes (see Joint Research Centre-European Commission 2008; see Saisana et al. 2005; Becker et al. 2017; Saltelli 2007).

The relevant problem here is that we need to consider different specifications of our composite indicator. Since a composite indicator is open to various possible specifications, queries concerning the final measure of interest can be made. Therefore, considering an equal weight is only an option that should be considered. Moreover, the choice of a single specification is typically subjective (Greco et al. 2019; Stefana et al. 2021). So in this respect, a different approach was proposed by Drago (2017). In this context, the idea is to obtain a comprehensive composite indicator using a Monte-Carlo simulation by applying different parameterizations of the initial specification (see Drago and Gatto 2022a, b; Drago 2021; Drago and Gatto 2019; Gatto and Drago 2021 for a different approach based on the ranks). For instance, it is possible to simulate the different weighting structures of the composite indicator. This approach also allows for evaluating the robustness of the initial results (see Gatto and Drago 2020). In this sense, it is usually said that the interval-based composite indicators tend to endogenize the sensitivity analysis typically performed on constructing a composite indicator (Gatto and Drago 2020). In this work, the innovation is related to using a different approach than intervals in constructing the final indicator. Here the final indicator is constructed utilizing a quantile indicator. Here the different descriptors are relevant quantiles obtained from the original indicator.

In this respect, the quantiles allow for an alternative representation to intervals (see for a different approach Davino et al. 2018). In this sense, the quantiles can show the entire data structure. They can allow us to understand better the structure of the Monte-Carlo simulations obtained if necessary (see the relevance of the understanding of the underlying simulations on the construction of interval-based composite indicators Drago 2019). In particular, their use can be interesting when we want to observe and compare the single quantiles in the different simulations and not explicitly an interval center. In this respect, the median of the indicator in their use is equivalent to the center computed using the intervals.

The two approaches are complementary because they are returned and can give different relevant information. In order to construct the composite indicator based on quantiles, we consider by simulation different specifications of the composite indicator and a single weighting of the composite indicator. In the end, we compute 2000 different simulations of the composite indicator. This result allows the construction of the final indicator based on the quantiles, computed considering all the outcomes obtained by simulation. The second composite indicator is based on the interval computed from the original simulation and the center and radius computed from the procedure. The center is the final value of the composite indicator, and it is typically used to compare the different regions (excluding extreme evaluations, which can be analyzed considering the lower bounds), where the range between upper and lower bound represents the variability of the different measurements.

3   Results

In order to interpret the results, we need to take into account the two different representations. In particular, we obtain two different tables to compare the results for the two composite indicators in Tables 1 and 2. We have considered the quantile and interval-based composite indicators approach to compare the different results.

We can observe higher ranks for Sardinia, Lombardy, Lazio, and Veneto from the quantile approach. First, we consider the median as the central value of the Monte-Carlo simulations. The result means a higher interest for the different queries in general, so there is a higher interest in entrepreneurial activities as financing a business in these regions. This result can also be due to the regional support for entrepreneurial activities. Interestingly, Veneto shows a substantial difference in their results considering each quantile. For instance, in the 10th quantile, Veneto performs worst than Lazio and Piedmont.

The result means a higher variability in the different results considering the different queries. So, Veneto performs better on some queries (showing attention to some approaches to women entrepreneurship in the region) than others. On the other hand, the Lazio, Lombardy, and Sardinia results are more consistent with their positions. That means that there is less variability in the different queries. In order to compare Table 1 with Table 2, the quantile-based composite indicator with the

Table 1   Quantile approach: ranking by 50th quantile

Region	q10	q25	q50	q75	q90
Sardinia	0.17	0.68	1.3	1.96	2.46
Lombardy	–0.07	0.11	0.86	1.69	2.38
Lazio	–0.04	0.23	0.58	1.19	1.69
Veneto	–0.23	0.05	0.45	1.33	2.08
Piedmont	–0.17	–0.03	0.2	0.65	1.04
Emilia-Romagna	–0.33	–0.13	0.19	0.54	0.87
Tuscany	–0.15	–0.01	0.13	0.3	0.42
Sicily	–0.17	–0.07	0.13	0.42	0.69
Apulia	–0.41	–0.25	–0.09	0.13	0.29
Campania	–0.41	–0.28	–0.15	0.2	0.47
Calabria	–0.54	–0.45	–0.31	–0.21	–0.05
Abruzzo	–0.63	–0.52	–0.36	0.31	0.8
Basilicata	–0.64	–0.53	–0.39	–0.16	0.07
Liguria	–0.65	–0.55	–0.44	–0.33	–0.24
Friuli-Venezia Giulia	–0.65	–0.57	–0.46	–0.36	–0.28
Marche	–0.67	–0.59	–0.49	–0.4	–0.32
Aosta	–0.9	–0.74	–0.58	–0.45	–0.34
Molise	–0.9	–0.74	–0.58	–0.45	–0.34
Trentino-Alto Adige/South Tyrol	–0.9	–0.74	–0.58	–0.45	–0.34
Umbria	–0.9	–0.74	–0.58	–0.45	–0.34

Source: Authors

Table 2 Ranking by the center of the interval-based composite indicator

Region	Lower bound	Center	Upper bound
Lombardy	–0.16	1.32	2.81
Sardinia	–0.1	1.31	2.72
Veneto	–0.37	1.05	2.48
Lazio	–0.15	0.9	1.96
Piedmont	–0.23	0.52	1.26
Emilia-Romagna	–0.41	0.32	1.04
Sicily	–0.23	0.29	0.81
Abruzzo	–0.69	0.17	1.04
Tuscany	–0.21	0.13	0.47
Campania	–0.47	0.07	0.6
Apulia	–0.47	–0.05	0.37
Basilicata	–0.69	–0.26	0.17
Calabria	–0.6	–0.3	0.01
Liguria	–0.7	–0.45	–0.2
Friuli-Venezia Giulia	–0.7	–0.48	–0.26
Marche	–0.71	–0.49	–0.28
Aosta	–0.97	–0.63	–0.29
Molise	–0.97	–0.63	–0.29
Trentino-Alto Adige/South Tyrol	–0.97	–0.63	–0.29
Umbria	–0.97	–0.63	–0.29

Source: Authors

interval-based composite indicator, we observe the general robustness of the different results obtained. It is relevant to note that the 0.01 and 0.99 quantiles in the intervals are used to evaluate the more extreme values of the composite indicator based on the different specifications sampled. Considering these results and interpreting them makes it possible to observe economic activity or interest in financing business, regulations, or economic opportunities.

Furthermore, it is possible to note from the results that where there is a higher volume of searches from public institutions, there is a higher volume of searches. That means the role of institutions in spreading the relevant information on women's entrepreneurship is very relevant. Also conceivable is the differentiation and evaluation of the many gaps between multiple organizations and institutions. In order to close these gaps, one must first identify what can be accomplished through effective communication tactics to raise awareness of the entrepreneurial potential of women (see, for instance, Hunt and Aghazamani 2017). Overall, better institutional communication can improve the interest and awareness of empowering women entrepreneurs. At the same time, the role of social media and good institutional communication cannot be under-evaluated.

In this sense, communication that uses different media channels can better empower women. Following Dewi (2020), evidence exists that social media may empower women in political, social, cultural, and economic arenas worldwide,

expanding in quantity and quality of evidence. In addition, the Internet and social media have a relevant role in collecting and exchanging information throughout the entrepreneurial process.

Social media has become an essential source of inspiration, empowerment, and networking for women entrepreneurs and professionals (Secundo et al. 2020; Khan et al. 2021). Users can share information, participate in debates, and communicate via social media platforms (Dewi 2020). The final effects of good communications and proper channels empower women's entrepreneurial process. Finally, it is also essential to consider the role of a clear and well-written text in the diffusion of information. This fact shows the difficulty of the process of institutional communication.

It is also interesting considering the worst results are for Aosta, Molise, Trentino-Alto Adige, and Umbria. Here, the result can be explained by a lower attraction of results in Google for the considered topics in these regions. It may result from women's entrepreneurial activities differing from searching for information online or less interest than in other regions.

The use of the quantiles is essential to identify some regions performing better on some specifications than others. The usefulness of the quantiles can become clear if we consider that some regions can perform better than others (as in the case of Veneto and Abruzzo). Abruzzo is an interesting case because it shows how the region is highly focused on specific queries. Generally, there is a bad performance on the indicator showing lower interest.

The interval-based composite indicator considers a different approach based on different quantiles. In this sense, the quantiles considered emphasizing the role of the extreme scenarios. These scenarios are helpful in economic policy where there is a difference in these rankings, meaning higher vulnerability or higher capacity to perform better considering specific specifications.

The results related to the interval-based composite indicator show strong robustness. The results seem consistent as rankings considering both the interval-based and the quantile-based composed indicator. The results are consistent in both methodologies, which means consistency in the evaluations.

4 Conclusions

We have considered a new approach for measuring women's entrepreneurship and economic activity in this work. Google measures the interest queries to track the interest in 2021 for the Italian regions. Furthermore, where research papers have shown the use of Google Trends to predict upcoming events successfully, it could be helpful to use this tool to track the economic activity of women's entrepreneurship interest.

The construction of a composite indicator is based on quantiles (a quantile-based composite indicator in which we also consider the interval between quantiles) and on intervals (interval-based composite indicator). The results show that four regions

perform better: Sardinia, Lombardy, Lazio, and Veneto. Furthermore, these regions show a higher interest in the considered queries, so in general, an higher interest in women's entrepreneurship. The data gathered shows that the number of searches is higher in areas with a greater level of activity by public entities disseminating information in this context. This suggests that the role of institutions in spreading pertinent information about women's entrepreneurship to the public may be of great importance. The limitations of this work are related to the fact that we have used only Google Trends as a source. In this respect, it is possible to consider more approaches to measure the interest in women's entrepreneurship. Future extensions can be devoted to exploring more aspects of women's entrepreneurship, which can be explored both considering web and social data.

Appendix: Queries Performed

"Imprenditoria femminile," "Impresa donna," "Fondo impresa donna," "Imprenditoria femminile 2021," "Bando imprenditoria femminile 2021," "Finanziamenti imprenditoria femminile," "agevolazioni imprenditoria femminile"

References

Becker, W., Saisana, M., Paruolo, P., & Vandecasteele, I. (2017). Weights and importance in composite indicators: Closing the gap. *Ecological indicators*, 80, 12–22.

Bianco, M. E., Lombe, M., & Bolis, M. (2017). Challenging gender norms and practices through women's entrepreneurship. *International Journal of Gender and Entrepreneurship*. https://www.emerald.com/insight/content/doi/10.1108/IJGE-10-2017-0060/full/html

Cardella, G. M., Hernández-Sánchez, B. R., & Sánchez-García, J. C. (2020). Women Entrepreneurship: A Systematic Review to Outline the Boundaries of Scientific Literature. *Frontiers in Psychology*, *11*, 1557.

Carranza, E.; Dhakal, C.; Love, I.. (2018). Female Entrepreneurs : How and Why Are They Different?. Jobs Working Paper;No. 20. World Bank, Washington, DC. © World Bank. https://openknowledge.worldbank.org/handle/10986/31004 License: CC BY 3.0 IGO.

Choi, H., & Varian, H. (2012). Predicting the Present with Google Trends. *The Economic Record*, *88*, 2–9.

Cockcroft, S., & Russell, M. (2018). Big data opportunities for accounting and finance practice and research. *Australian Accounting Review*, 28(3), 323–333.

Dal Mas, F., Paoloni, P., & Lombardi, R. (2019). Wellbeing of women entrepreneurs and relational capital: A case study in Italy. In *The Wellbeing of Women in Entrepreneurship* (pp. 232–243). Routledge.

Davino, C., Dolce, P., Taralli, S., & Vinzi, V. E. (2018). A quantile composite-indicator approach for the measurement of equitable and sustainable well-being: A case study of the Italian provinces. *Social Indicators Research*, *136*(3), 999–1029.

Demartini, P., and Paoloni (2011) "Assessing human capital in knowledge intensive business services." Measuring Business Excellence. Bradford Vol. 24, Fasc. 1, (2020): 114–130. DOI: https://doi.org/10.1108/MBE-08-2019-0086

Dewi, A. S. (2020). Investigating the Impact of Instagram on Women Entrepreneurs' Empowerment. In Proceedings of the Annual International Conference on Social Sciences and Humanities (AICOSH 2020) (pp. 58–61).

Donadelli, M., & Gerotto, L. (2019). Non-macro-based Google searches, uncertainty, and real economic activity. *Research in International Business and Finance,* 48, 111–142.

Drago, C. (2017), Interval Based Composite Indicators. FEEM Working Paper No. 42.2017, Available at SSRN https://ssrn.com/abstract=3038751

Drago, C.(2019) Decomposition of the Interval Based Composite Indicators by Means of Biclustering. CLADAG 2019 12-th Scientific Meeting Classification and Data Analysis Group Cassino, September 11–13, 2019 Book of Short Papers/Abstracts , Available at SSRN: https://ssrn.com/abstract=3568357

Drago, C. (2021). The Analysis and the Measurement of Poverty: An Interval-Based Composite Indicator Approach. In *Economies* (Vol. 9, Issue 4, p. 145). https://doi.org/10.3390/economies9040145

Drago, C., & Gatto, A. (2019). A Robust Approach to Composite Indicators Exploiting Interval Data: The Interval-Valued Global Gender Gap Index (IGGGI). In *Advances in Gender and Cultural Research in Business and Economics* (pp. 103–114). https://doi.org/10.1007/978-3-030-00335-7_7

Drago, C., & Gatto, A. (2022a). An interval-valued composite indicator for energy efficiency and green entrepreneurship. *Business Strategy and the Environment 31*(6): 2107-2126

Drago, C., & Gatto, A. (2022b). Policy, regulation effectiveness, and sustainability in the energy sector: A worldwide interval-based composite indicator. *Energy Policy, 167*, 112889.

Gatto, A., & Drago, C. (2020). Measuring and modeling energy resilience. *Ecological Economics: The Journal of the International Society for Ecological Economics.* https://www.sciencedirect.com/science/article/pii/S0921800919301570

Gatto, A., & Drago, C. (2021). When renewable energy, empowerment, and entrepreneurship connect: Measuring energy policy effectiveness in 230 countries. Energy Research & Social Science, 78, 101977.

Gómez Martínez, R., Prado Román, M., & Mercado Idoeta, C. (2015). Google Search Activity as Entrepreneurship Thermometer. In M. Peris-Ortiz & J.-M. Sahut (Eds.), *New Challenges in Entrepreneurship and Finance: Examining the Prospects for Sustainable Business Development, Performance, Innovation, and Economic Growth* (pp. 225–233). Springer International Publishing.

Google Trends (2022) https://www.google.com/trends

Greco, S., Ishizaka, A., Tasiou, M., & Torrisi, G. (2019). Sigma-Mu efficiency analysis: A methodology for evaluating units through composite indicators. *European Journal of Operational Research,* 278(3), 942–960.

Hunt, C. A., Aghazamani, Y. (2017). Empowerment in Tourism: A Review of Peer-Reviewed Literature. *Tourism Review International*, 4(21), 333–346.

Joint Research Centre-European Commission. (2008). Handbook on constructing composite indicators: methodology and user guide. OECD publishing.

Khan, R. U., Salamzadeh, Y., Shah, S. Z. A., & Hussain, M. (2021). Factors affecting women entrepreneurs' success: a study of small- and medium-sized enterprises in emerging market of Pakistan. *Journal of Innovation and Entrepreneurship*, 10(1), 11.

Matias, Y. (2013, October). Nowcasting with Google Trends. In International Symposium on String Processing and Information Retrieval (pp. 4–4). Springer, Cham.

Mayer, H. (2008). Segmentation and segregation patterns of women-owned high-tech firms in four metropolitan regions in the United States. *Regional Studies*, 42(10), 1357–1383.

Minniti, M. (2010). Female Entrepreneurship and Economic Activity. *The European Journal of Development Research*, 22(3), 294–312.

Moletta, J., de Carvalho, G. D. G., do Nascimento, R. F., Barboza, B. M. L., Resende, L. M., & Pontes, J. (2021). Business networks of women entrepreneurs: an analysis of the expectation and reality of factors that affect trust in a business network. *Journal of Intelligent Manufacturing*, 1–16.

Nagao, S., Takeda, F., & Tanaka, R. (2019). Nowcasting of the US unemployment rate using Google Trends. *Finance Research Letters*, 30, 103–109.

Paoloni, P., & Demartini, P. (2016). Women in management: perspectives on a decade of research (2005–2015). *Palgrave Communications*, *2*(1), 1–7.

Paoloni, P., & Demartini, P. (2012). The relational capital in female smes. *Journal of Academy of Business and Economics*, 12(1), 23–32.

Paoloni, P., & Dumay, J. (2015). The relational capital of micro-enterprises run by women: the startup phase. *Vine*.

Saisana, M., Saltelli, A., & Tarantola, S. (2005). Uncertainty and sensitivity analysis techniques as tools for the quality assessment of composite indicators. *Journal of the Royal Statistical Society*: Series A (Statistics in Society), 168(2), 307–323.

Saltelli, A. (2007). Composite indicators between analysis and advocacy. *Social indicators research*, 81(1), 65–77.

Secundo, G., Del Vecchio, P., & Mele, G. (2020). Social media for entrepreneurship: myth or reality? A structured literature review and a future research agenda. *International Journal of Entrepreneurial Behavior & Research*.

Sharma, S. (2018). Emerging dimensions of women entrepreneurship: Developments & obstructions. *Economic Affairs*, *63*(2), 337–346.

Stefana, E., Marciano, F., Rossi, D., Cocca, P., & Tomasoni, G. (2021). Composite Indicators to Measure Quality of Working Life in Europe: A Systematic Review. *Social Indicators Research*, 157(3), 1047–1078.

Wang, Q. (2013). Industrial Concentration of Ethnic Minority- And Women-Owned Businesses in the United States. SSRN Journal. https://doi.org/10.2139/ssrn.2286372