LaSEI
School of Economic Intelligence & International Relations
Escuela de Inteligencia Económica y RRII
Universidad Autónoma de Madrid

01/2021

Vol. 3, n 1, pp. 1 - 9
ISSN: 2659-9805
Editor in-Chief: Javier Oubiña

JOURNAL OF ECONOMIC & BUSINESS INTELLIGENCE

Universidad Autónoma de Madrid UAM

TAILORING BUSINESS INTERNATIONALIZATION SUPPORT BY MEANS OF ANALYTICS

Joaquín María Núñez Varo*

Abstract

At ICEX, evaluations and analytics have enabled us to develop a system capable of revealing with great precision any company's needs as well as its evolution in competitiveness, creation of added value and employment. The system can actually prescribe the best support to be offered according to the company's capabilities, international history and the effectiveness of the support in overcoming the market barriers for that particular company. And this is just the beginning.

Key words: Analytics, artificial intelligence, business internationalisation, evaluation of public policies, public support for business internationalisation, recommender systems

Resumen

Los resultados de las evaluaciones de ICEX así como el análisis de los datos disponibles en la institución han permitido poner a punto un sistema capaz de revelar con gran precisión las necesidades de una empresa, su evolución en términos de mejora de la competitividad, de la creación de valor añadido y de empleo, así como de prescribir el mejor servicio que podemos ofrecerle en función de sus capacidades, su historial internacional y la efectividad del propio servicio para superar sus barreras en el mercado que quiera acometer. Y este es solo el comienzo.

Palabras clave: Analítica, apoyo público a la internacionalización empresarial, evaluación de políticas públicas, inteligencia artificial, internacionalización empresarial, sistemas de recomendación

* Analyst of the Evaluation Department at ICEX Spain Export and Investment

1. Introduction

ICEX's Evaluation Unit was created in 2011 with the mandate of finding out the effectiveness of the different kinds of support in the agency portfolio. From the beginning, our objective was for evaluations to actively inform the ongoing improvement of our support for the internationalisation of businesses. And this goal could only be achieved by enriching classic evaluation methodologies with analytics.

Our evaluation strategy consists of three stages:

First stage: establishment of the operational models for the different business support programmes.

Second stage: identification of control groups for each type of company and variables in administrative records representing those aspects of global competitiveness where our first evaluation stage has shown an effect on the company after our support. This phase is now in the making, although by means of feature engineering we have already uncovered relevant topological invariants, i.e. coherent metrics that accurately inform the evolution of a company in terms of competitiveness.

Third stage: we propose that the best indicator of the long-term impact of public policies in support for business internationalisation is the return of public investment via taxation associated with those company results that can be safely attributed to our support through evaluation techniques. We are aware that ICEX cannot carry out this stage alone, but we do think that we should be part of the research team that undertakes it.

As a result of the evaluation work carried out in the first and second stages, impact evaluations have revealed effects one, two and three years after receiving the main types of ICEX support on 12 competitiveness variables, while needs assessments have revealed both each company's real drivers for competitiveness in the global market and its interest and price intention in 89 kinds of support designed to meet their different needs. The results are differentiated by type of company.

All the work done to date can be seen in detail in Núñez (2017) and Núñez (2020). From now on, we are going to focus on the analytical strategy we have followed.

For each kind of support, we have carried out a maximal segmentation by means of *ad hoc* clustering based on survey responses and certain variables coming from administrative records. Then we have constructed classifiers to predict the cluster to which any given company belongs. So far, all the classifiers we have built have an accuracy between 79% and 99.995%, with an area under the ROC (Receiver Operating Characteristic) curve greater than 0.96 (which implies a great deal of cleanliness in the classification, meaning that we cannot improve the accuracy of the classification with

Table 1

Intervention logic for ICEX generic support and relevant questions

Entity	Variables	Relevant questions
Beneficiary	Company's characteristics	Who are they?
Treatment	ICEX's support	What does it consist in? How do companies use our support? Do they follow any strategy?
Output	Service or product delivered *(Quality and satisfaction surveys)*	What are we offering? What do they need?
Outcomes	Improvement in competitiveness *(greater added value of the Spanish offer)* Exports Foreign investment Profit and loss account Quality employment	How do we measure them? Are outcome indicators defined per kind of support? How are these indicators linked to the performance indicators reflected in the strategic plan of the Secretariat of State for Trade?
Economic outcomes	Return on investment via taxes	How much wealth is leveraged by the allocated resources?

Source: ICEX Evaluation Department.

the available information). It should be noted that the most relevant variables are those captured by our evaluation surveys, since, when these variables are discarded, accuracy drops to 54% in the best-case scenario: the misclassification rate rises to a minimum of 46%!

Using sequential analysis techniques, we can predict with great precision the most likely kind of support a company will request from us the following two years, taking into account the sequence of support instruments requested in the past. We can also know the most likely gateway to our wide portfolio of services for companies that have never knocked at our door before according to their characteristics. Furthermore, using the evidence of our impact evaluations, it is straightforward to calculate which support best fits the company's internationalisation needs, and we can provide this particular recommendation in their managers' own language so that they can understand not only the recommendation but also the explanation of why it meets their company's needs. We use diffuse logic for this calculation, that is, all possible paths all followed together with their truth probabilities and the results of the different paths are subjected to an interference process from which coherent results come out reinforced while incoherent results just cancel out. For the time being, we use plain Kolmogorov's axioms for our probability models, but we are considering the use of Växjö's contextual probability model, developed by the Russian mathematician Andrei Khrennikov as from 2000 as a generalisation of the quantum probability model by allowing probability amplitudes to take values in a two-dimensional Clifford algebra (the algebra of hyperbolic complex numbers). This probability model overcomes the problem of the systematic violation of Savage's sure-thing principle, which states that the decision of an economic agent does not depend on their knowledge of the occurrence of a certain event. Violation of this principle implies a lack of rationality in human decisions, particularly economic ones, and is therefore applicable to our field of study.

Table 2

Some of our classifiers

	Classifier 1	Classifier 2	Classifier 3	Classifier 4
Type	Needs	Needs	Effectiveness	General
Classification model	Logistic regression	Gradient boosting	Logistic regression	Logistic regression
Objective	Spanish Official Pavilions at Trade Fairs	Support to commercial prospection	Tailor-made services	Company profiling
Represented population	Companies using support for participation in Spanish Official Pavilions at Trade Fairs	Companies using support services	Companies using tailor-made services	Companies using ICEX's services
Variables	7	11	7	65
Dimensions	3	4	3	22
Precision	99%	79%	93%	79%
Cleanliness	1.000	0.973	0.991	0.995

Source: ICEX Evaluation Department.

We include in our needs assessment surveys a diagnosis block, from which we can infer the company's competitiveness drivers by means of applying a common technique to extract preferences in different items as used in recommender systems, which avoid the problem of different perception thresholds for each company, that is, the divergence in the degrees of demand and reference between companies depending on their different state of maturity and development. Specifically, we standardise the responses of each company according to the distribution of all their responses in all the diagnostic variables and we subject these new standardised variables to a dimensional scaling analysis. As a result, we have identified three axes, which we call opposing factors, and the position along these axes classifies each company in terms of its real capabilities. Each of these axes, which we detail below, is in turn made up of two different competing dimensions.

1. Axis 1 shows the balance between the following dimensions:
    1.1. International culture and networking
    1.2. Human and financial resources.

Table 3

Competitiveness drivers for the different types of companies requesting support for their own prospect initiatives

	International culture and networking	Human and financial resources	Continuous training and analytical capacity	Traditional tasks of a commercial department	Financial resources and productive capacity	Talent of the commercial human resources
Type 1						
Type 2						
Type 3						
Type 4						
Type 5						
Type 6						
Type 7						
Type 8						

The company stands on this diagnostic dimension as a competitive factor

This diagnostic dimension is balanced as a competitiveness factor

The company does not stand on this diagnostic dimension as a competitive factor

Types 6, 7 and 8 are different kinds of microenterprises.
Source: ICEX Evaluation Department.

2. Axis 2 shows the balance between the following dimensions:
   2.1. Continuous training and analytical capacity.
   2.2. Traditional tasks of a commercial department.
3. Axis 3 shows the balance between the following dimensions:
   3.1. Financial resources and productive capacity.
   3.2. Talent of the commercial human resources.

We call competitiveness driver the dimension within each axis that is reinforced in each type of company. Table 3 shows the competitiveness drivers of the eight different types of company that emerged from our 2018 needs assessment to improve our support to companies' own initiatives to prospect for opportunities in the international market.

2. Company profiling

A few years ago we started work on the creation of a synthetic index to assess the internationalisation potential of a company. A first study was carried out in 2017 to draw the profiles of companies that had used ICEX support services between 2009 and 2016, taking into account 65 variables including both characteristics and competitiveness performance indicators which turned out to represent various dimensions of the 22 latent factors uncovered by a robust factor analysis. This study revealed 10 very different groups of companies, for which differentiated needs in terms of internationalisation can be inferred. At present, the profiling is being extended to companies having used ICEX services between 2003 and 2019 (which total over 180,000 entities among different corporate forms of companies and registered self-employed workers). This new study will determine a segmentation of companies according to their needs and internationalisation potential and will fine-tune our classifiers in order to spot companies to which we can offer those of our services which best meet their current or future needs. As part of this work, we first had to complete by means of multiple imputation algorithms the data collected from the Trade Register for 27 longitudinal variables measured as from 2008. In particular, we applied both the expectation-maximization method and the fully conditional specification imputation method (Markov chain-based Monte Carlo method, MCMC), which output very similar results, and then, as a final decision rule for the effective

Table 4

Quality of the imputation of the set of selected variables from the Trade Register for ICEX's clients since 2003

Validity criterion	Value
Maximum imputation error	3%
Maximum variation in individual and conditional entropy	1%
Maximum variation in mutual information between non-redundant main components	1%

Source: ICEX Evaluation Department.

Table 5

Types of companies by export flows

Type (exports)	Count	Years	Mean	Median	Minimum	Maximum	Deviation	Range
TODAS	932 534	2.41	304 426.65	279 197.18	99 479.39	609 536.74	518 310.87	268 650.95
1	45 425	5.05	5 799.59	4 406.64	1 319.25	12 713.44	5 203.69	9 200.10
2	22 101	7.31	4 515 388.51	4 421 721.76	2 430 204.37	6 928 132.09	1 498 617.52	2 523 362.58
3	68 352	9.28	2 143 513.37	2 013 920.51	300 886.86	4 645 824.82	1 468 697.32	2 340 370.26
4	46 977	2.00	88 372.36	88 372.36	33 099.28	143 645.43	78 167.93	110 546.14
5	70 306	2.00	904.24	904.24	420.21	1 388.27	684.52	968.06
6	26 697	8.98	628 758.95	159 354.11	20 544.25	2 771 171.39	950 262.30	1 089 025.45
7	34 344	3.90	764.37	639.05	274.83	1 429.60	559.47	1 033.96

Values in euros.
Data: Customs Data (Spanish Tax Office) between 2000 and 2018 by fictitious tax identification number
File size: 8 195 007 records (corresponding to the cross between fictitious NIF-Year-Country of exportation)
Total sequences: 932 534
Technique: K-means clustering
Source: Spanish Tax Office and ICEX Evaluation Department.

imputation of each specific missing value, we calculated a randomized linear combination between both methods using a uniformly-distributed parameter. Table 4 shows the imputation quality.

3. Profiling by export flows

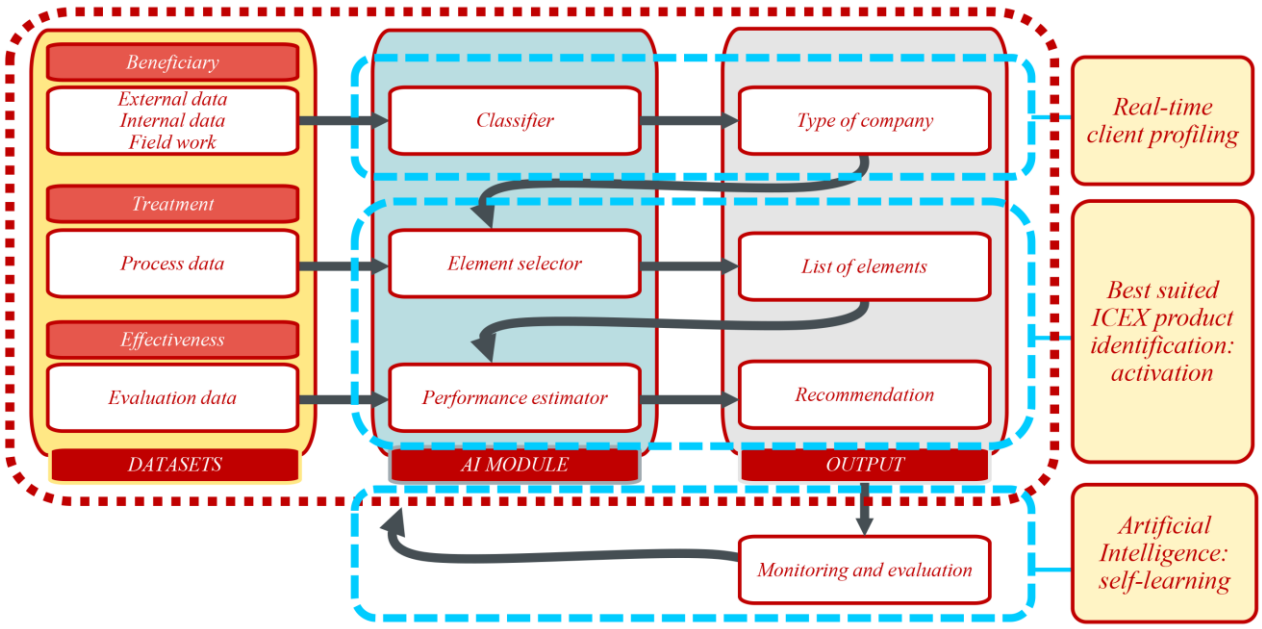
Besides, we have carried out a K-means clustering of export flows between 2000 and 2018 using the disaggregated, anonymised data from the Spanish Tax Office, which has output 7 very different types of company in terms of the distribution of export volume over time, which is tantamount to a company's cash flow. An advantage of this approach lies in its taking into account as fundamental elements of these export flows the stability over time as well as the survival of the company in the international market. Table 5 shows the descriptive statistics of the 7 different resulting types. We have also subjected this comprehensive over-8-million-register database to sequential analysis techniques by geographical area, which, as in the case of its application to the use of our kinds of support by companies, allows us to calculate with great precision the probability that a company can enter a new geographical market or simply remain in it depending on the areas it has previously traded in and the type of company it is.

4. Self-diagnosis and recommender system based on Artificial Intelligence

All this knowledge has encouraged us to design a self-diagnosis and recommender system that can prescribe each of our clients the most suitable kind of support in terms of both their effectiveness and the added value of public resources. Feeding the system with data of the real effectiveness of our support will enable self-adjustment –learning–, so it will enter the paradigm called Artificial Intelligence. Graph 1 shows the blueprint of the designed system, in whose implementation we are now deeply committed and involved.

Graph 1

Blueprint for ICEX self-diagnosis and recommender system based on Artificial Intelligence



Source: ICEX Evaluation Department.

The exceptional situation brought about by the COVID-19 pandemic has provided an opportunity to test both the accuracy and the reliability of the system. ICEX carried out last spring a we-are-by-your-side campaign for Spanish companies, with the dual objective of listening the unscreened reality of each company in the face of the upbrought situation and to be able to identify the most appropriate kind of support for them. In particular, we compared the kind of support which resulted in the campaign, identified by means of a phone conversation with a human agent belonging to one of the different Territorial and Provincial Trade Departments throughout Spain, and the recommendation extracted from our self-diagnosis and recommender system. The rate of agreement in the common sample (that is, for those records with a kind of support identified in the campaign for which there is a recommendation from the system) scores 91.24%. Where there was not full agreement, both the recommendation and the diagnostic of our system were considerably more detailed than the lead detected by a human agent. We must highlight that the importance of this critical field test does not really lie in the rather high accuracy rate, but in the fact that the entire system was built without taking into account as a factor as extraordinary a crisis as the one caused by the outbreak of the pandemic in the business environment, while the identification of the kind of support in the human-interface campaign did indeed take the exceptional situation into account. The high accuracy rate and the greater detail of the output of the self-diagnosis and recommender system is a proof of its robustness in the face of such drastic changes as the one we have experienced, the reason being its taking into consideration not a one-off situation of the company at any determined time but its entire evolution and growth prospects using as a reference the rest of mirror companies.

5. User experience gamification in the new ICEX portal system

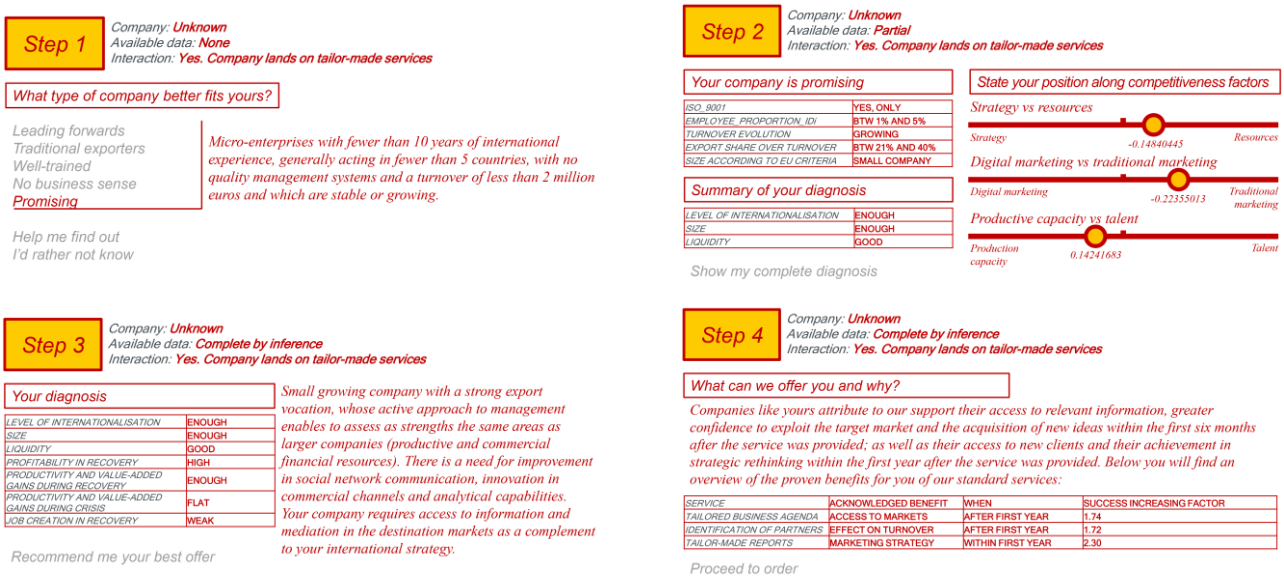
The self-diagnosis and recommender system itself has opened up the possibility of guiding navigation in our portal system through user experience gamification in order to lead them to the product or service that best meets their needs. The process is essentially guided by the following relationship:

Product = F (Needs, Effectiveness, Availability)

In other words, the product to be offered a company depends on the company's needs, the effectiveness of the product on the company itself and its availability in our portfolio and activity plan. Table 6 shows the maximum number of questions required to determine the most appropriate product or service for a company. Although this is a

Graph 2

Concept proposal for the user experience gamification for an unknown company interacting within the ICEX portal system



Source: ICEX Evaluation Department.

limited and quite acceptable number, the actual number of questions will normally be smaller, as there is no point in asking a company for information that we already know.

We will show with a couple of examples how the process would unfold, their concept proposals being shown in graphs 2 and 3. Imagine you land on a given ICEX website and you work in a company unknown to us, that is, a company not yet registered in our customer base. We start from the fact that you are interested in let's say the webpage for ICEX tailor-made services, which triggers a specific context. In the first place, we would ask you the type of company you work at from a series of tags embodying marketing-oriented labels for companies along with a precise description of the type of company so that you can locate that which best portrays your company. Once the type of company has been selected, we are able to determine the value of 7 characterisation variables with a minimum precision of 80%, although depending on the search context (determined by the type of service you are interested in) a 99.995% precision can be actually reached. On the following screen we would show you these characteristics as well as a summary of your diagnosis, consisting of a selection of the 22 latent factors depicting the evolution of your company's competitiveness from 2008 to date. We would also show you a block of competitiveness drivers consisting of three regulators that you can slide at will until they mirror the actual situation of your company. All this information calls for a much more accurate profiling process, since all the different context-specific classifiers come into operation in a diffuse way in order to offer your complete diagnosis on the next screen. This diagnosis shows a more extensive selection of the 22 latent factors, which the company can also modify at will, along with a text informing you about the international competitiveness potential of

Table 6

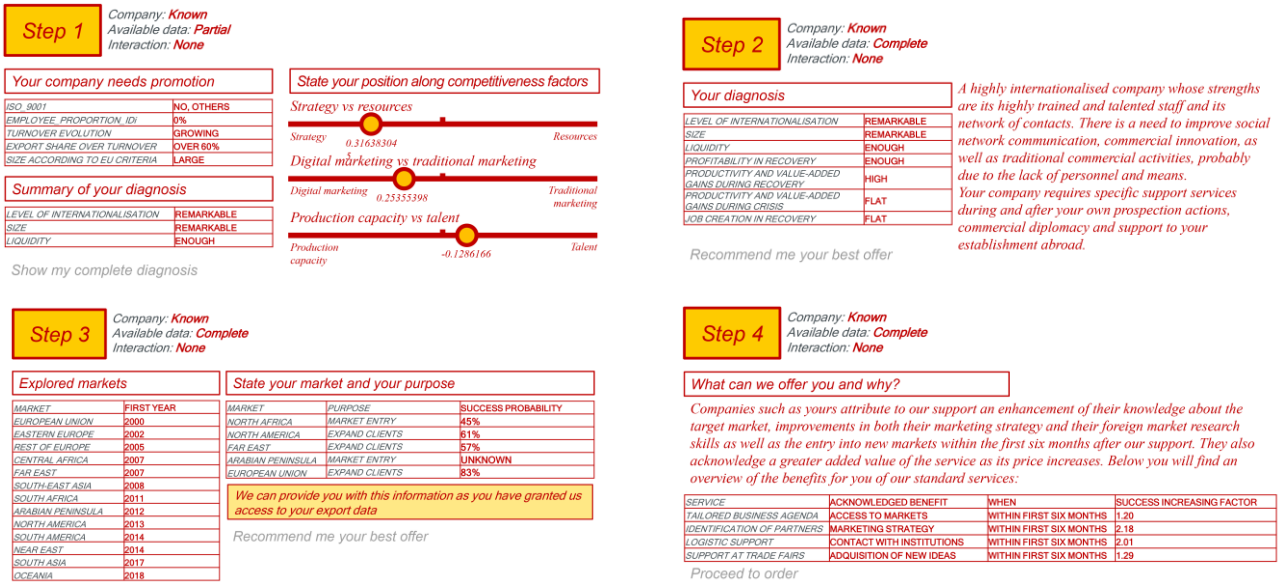
Number of bits (binary Yes/No questions) needed to determine the most appropriate product for a company

	Expression	Needs (1)	Markets (2)	Kind of support (3)
Number of possible values	n	8	21	8
Bits	$log_2 n$	3	4,39	3
Necessary questions	$Ceiling(log_2 n)$	3	5	3

Source: ICEX Evaluation Department.

Graph 3

Concept proposal for the user experience gamification for a known company after accessing its personal area at ICEX portal system



Source: ICEX Evaluation Department.

your company and what is required to optimise its internationalisation process. Eventually, we can display at your choice those of our services that best meet your needs and context.

Imagine now that you work at a known company, that is, a company already registered in our customer base, and you access the home page or your personal area without any other interaction. As we know both you and your company, we would show the kind of company straight away as well as the regulators that you can slide at will until they mirror the actual situation of your company. Again, we offer you an updated profile along with a text informing you about the updated international competitiveness potential of your company and what is required to optimise its internationalisation process. Next, if someone in your company has allowed us to have the detailed data on exports of goods from the Spanish Tax Office, we would show you on a screen the geographical markets your company had operated in along with the first year of operation. We would also show you an interactive table where you can select different markets of your interest and the objective you are pursuing so that the system can return the probability of success in achieving that objective in that market. Eventually, we can display at your choice those of our services that best meet its needs and context. We can even fine-tune these services depending on the geographical market you are targeting.

6. References

Aggarwal (2016). Recommender Systems The Textbook. Aggarwal, Charu C. (Ed. Springer, 2016) https://link.springer.com/content/pdf/10.1007%2F978-3-319-29659-3.pdf

De Lucio (2018). Capítulo sobre necesidades de las empresas en su proceso de internacionalización de la obra colectiva "La internacionalización de la empresa española: Situación y propuestas". Paula de Diego, Juan de Lucio, Maribel Martínez, Joaquín Núñez (Ed. Civitas, 2018). https://www.thomsonreuters.es/es/tienda/duo-papel-ebook/la-internacionalizacion-de-la-empresa-espanola--situacion-y-propuestas-duo/p/10012891

Estrategia de Internacionalización, S. G. de. (2019). El Plan de Acción para la Internacionalización de la Economía Española 2019-2020. *Boletín Económico de ICE*, (3111). https://doi.org/10.32796/bice.2019.3111.6827

Estudios y Evaluación de Instrumentos de Política Comercial, S. G. (2020). Informe de Evaluación del Plan de Acción para la Internacionalización de la Economía Española 2017-2018 https://comercio.gob.es/es-es/estrategia_internacionalizacion/Paginas/informe-evaluacion-plan-2017_18.aspx

HM Treasury (2011). HM Treasury Magenta book. Guidance for evaluation. (Ed. HM Treasury, 2011)

https://www.gov.uk/government/publications/the-magenta-book

Kumar (2019). Introduction to Data Mining (Second Edition). Tan, Pang-Ning; Steinbach, Michael; Karpatne, Anuj; Kumar, Vipin. (Ed. Pearson, 2019) https://www-users.cs.umn.edu/~kumar001/dmbook/index.php

NIESR (2015). Evaluating the impact of UKTI trade services on the performance of supported firms. Ana Rincón-Aznar, Rebecca Riley, Anna Rosso. (Ed. NIESR, septiembre 2015) https://www.gov.uk/government/publications/evaluating-the-impact-of-ukti-trade-services-on-the-performance-of-supported-firms

Núñez (2017). La evaluación del apoyo público a la internacionalización empresarial, Joaquín Núñez. (Disponible en la red Academia, mayo 2017) https://www.academia.edu/37555717/La_evaluaci%C3%B3n_del_apoyo_p%C3%BAblico_a_la_internacionalizaci%C3%B3n_empresarial

Núñez (2020, mayo). El valor estratégico de la evaluación en ICEX. Joaquín Núñez, Mar Gómez. Boletín Económico de ICE, 3123, 39-62 http://www.revistasice.com/index.php/BICE/article/view/7003/7015

Peña (2002). Análisis de datos multivariantes. Peña, Daniel. (Ed. McGraw Hill, 2002) https://www.mheducation.es/analisis-multivariante-de-datos-9788448136109-spain-group

Pérez (2015). Técnicas de evaluación de impacto. Pérez López, César; Moral Arce, Ignacio. (Ed. Garceta, 2015)

UKTI (2014). Technical Note: UKTI's monitoring and evaluation evidence. (Ed. UKTI, junio 2014) https://www.gov.uk/government/publications/technical-note-uktis-monitoring-and-evaluation-evidence/technical-note-uktis-monitoring-and-evaluation-evidence

Word Bank (2016). Impact Evaluation in Practice (2nd. Edition). Gertler, Paul J.; Martinez, Sebastian; Premand. Patrick; Rawlings, Laura B.; Vermeersch, Christel M. J. (Ed. Word Bank, 2016) https://openknowledge.worldbank.org/handle/10986/25030