

Article

The Dynamics between Structural Conditions and Entrepreneurship in Europe: Feature Extraction and System GMM Approaches

Ana Borges ^{1,*}, Aldina Correia ^{1,†}, Eliana Costa e Silva ^{1,†} and Glória Carvalho ^{2,†}

¹ CIICESI, ESTG, Politécnico do Porto, 4610-156 Felgueiras, Portugal; aic@estg.ipp.pt (A.C.); eos@estg.ipp.pt (E.C.e.S.)

² ESTG, Politécnico do Porto and Universidade do Minho, 4704-553 Braga, Portugal; mgfc@estg.ipp.pt

* Correspondence: aib@estg.ipp.pt

† These authors contributed equally to this work.

Abstract: Structural conditions and population characteristics of countries have been identified in the literature as factors for an individual to become, or to have intentions of becoming, an entrepreneur. However, this is still a subject under research, which has become increasingly relevant and could be crucial in the current challenges of European countries. In this work, the factors for entrepreneurial intentions and entrepreneurship activity are studied. More precisely, the structural conditions of European countries, which has changed over the last two decades, is analysed. The aim is to describe this behaviour and to state the main conditions for developing entrepreneurship activities and the intentions to become an entrepreneur. To achieve this purpose, feature extraction, namely, principal component analysis and dynamic longitudinal approaches are used. In particular, we propose that the system-generalised method of moments (GMM) model is adequate in this situation. The results suggest that the structure of the European framework conditions for entrepreneurship, obtained using the Factor Analysis year by year, is quite diversified until 2008, while after 2008, it is more stable. Moreover, it is concluded that the conditions associated with entrepreneurial intentions and entrepreneurial activity differ between these two time periods. Hence, the dynamic aspect of the structural conditions that affect entrepreneurial activities or intentions should be acknowledged.

Keywords: entrepreneurship; European structural conditions; entrepreneurial intentions; entrepreneurial activity; longitudinal models; system GMM; factor analysis

MSC: 62H25; 62P20; 62M10



Citation: Borges, A.; Correia, A.; Costa e Silva, E.; Carvalho, G. The Dynamics between Structural Conditions and Entrepreneurship in Europe: Feature Extraction and System GMM Approaches.

Mathematics **2022**, *10*, 1349. <https://doi.org/10.3390/math10081349>

Academic Editor: Manuel Alberto M. Ferreira

Received: 8 March 2022

Accepted: 13 April 2022

Published: 18 April 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Entrepreneurship can be defined as a person's capacity to put their ideas into action. It requires the ability to articulate project planning, management, to take calculated risks and to innovate and be creative, in order to achieve previously defined goals [1]. In recent decades, entrepreneurship has gained increasing attention and becomes a crucial topic on both researchers' and political leaders' agendas [2]. From the perspective of political leaders, entrepreneurial activity is a source of innovation, competitiveness and economy development, while academia focuses on deepening the knowledge about this topic [2].

In an extensive meta-analysis [3], it is pointed out that the most common measure of entrepreneurship is the Total early-stage Entrepreneurial Activity (TEA), which is defined as the percentage of individuals aged 18–64 who are either starting new businesses, or are currently running a business for less than 42 months. Furthermore, TEA is attractive as a measure of entrepreneurship because it does not rely on official government statistics. This is an advantage in cross-country studies, as it facilitates standardisation, since the definition of entrepreneurship varies between governments.

Entrepreneurial Framework Conditions (EFC) and Entrepreneurship Intentions (EI) are also two important concepts associated with entrepreneurship.

EIs are defined, by [4], as the entrepreneurial orientation, vocational aspirations and the desire to own a business (Thompson, 2009), measured by the percentage of adult-age population (18–64 years old) who are latent entrepreneurs and who intend to start a business within three years.

The EFC, from the Global Entrepreneurship Monitor (GEM) research project, are defined by Valliere [5] as the environmental conditions that support entrepreneurial activity at the national level. Subsequently to Valliere's work, and until today, as illustrated by Orobia et al. [6], several authors focused on the study of countries' EFC.

GEM is a large-scale database for internationally comparative entrepreneurship and is funded in 1997. It includes information concerning many aspects of entrepreneurship activities of a large number of countries. This research project is the world's largest continuing investigation of entrepreneurial dynamics [7]. The project's initial report, released in 1999, covered ten developed economies, including eight from the OECD (Canada, Denmark, Finland, France, Germany, Israel, Italy and the United Kingdom), as well as Japan and the United States [7]. Ever since, it has expanded to cover a diverse range of economies from around the globe [8].

The GEM survey is based on the collection of primary data from at least 2000 randomly selected adults (18–64 years of age) in each economy via an annual Adult Population Survey (APS). It provides standardised data on entrepreneurial activities and attitudes. Furthermore, national teams collect experts' opinions on the components of the entrepreneurship ecosystem through a National Expert Survey (NES) [7] using standardised questionnaires. According to the 2019/2020 GEM Global Report, 50 economies participated in the GEM 2019 Adult Population Survey. Moreover, over 150,000 individuals participated in extended interviews. The GEM project is of great importance to the research on entrepreneurship, as with the data provided by GEM it is possible to understand the evolution of entrepreneurship through time and between countries. Furthermore, it allows us to pose research questions that could not have been addressed before [9].

Although several authors have researched the conditions for entrepreneurship all over the world, to the best of our knowledge, the study of the impact of the EFCs on the EI and on the TEA in Europe has not been systematically addressed. Therefore, the present study intends to contribute to the understanding of this emerging concept, as well as analyse which EFCs have an impact on Entrepreneurial Behaviour and Attitude (EBA), such as EI and TEA, using the two GEM databases above-mentioned in the European context, similarly to [10,11]. Specifically, by returning to statistical methods that allow a dynamic structure between subjects (in this case countries) and through time, we aim to answer the two following research questions:

- **Question 1**—How do the EFCs influence the EI?
- **Question 2**—How do the EFCs influence the TEA?

As previous studies have demonstrated a dynamic behaviour of the EFC indicators between countries and through time [12], the correlation structure of EFC indicators is analysed for the period between 2001 and 2019, in order to understand whether it remains the same or whether there are changes. For this purpose, FA, with the varimax approach, is used. Next, it is analysed how the EFCs relate to EI and TEA, through a longitudinal study of an unbalanced panel data set on 34 European countries between 2001 and 2019.

According to several studies on GEM data (e.g., [13,14]), entrepreneurship is heterogeneous across countries. Hence, it is important to take into account the intra-country specificities when conducting these types of analyses. For that, different specifications of longitudinal models are tested to understand which EFCs have a significant effect on the countries' EI or TEA over time. Additionally, similarly to the work of [15], the estimations are controlled for per-capita GDP, population growth and type of industry, as the country's wealth may have a significant impact on the nature of its entrepreneurial activity [16].

The GEM survey data has the advantage of being less susceptible to biases in coverage across regions [3]. In fact, due to less thorough data collection or economic and regulatory environments that discourage entrepreneurs from formally registering their activities, there is the risk that low-income countries are underrepresented in official statistics.

The rest of the paper is organised as follows. In Section 2, details on the data set used are presented, as well as the statistical methodologies. The results obtained when using the FA and system GMM are presented in Section 3. Finally, Section 4 presents the main conclusions and proposals for future work.

2. Data and Methodology

2.1. Data

As expressed in the previous Section, several researchers have recognised EI, TEA and EFCs as relevant indicators of entrepreneurial activity (e.g., [3–16]); thus, these will be used in the present study. More precisely, the data set used in this work contains data on 34 European countries, from 2001 to 2019, on the EBA, namely EI and TEA, and on the following 12 EFCs from GEM, as described in [8]:

- EFC1: Financing for entrepreneurs;
- EFC2: Governmental support and policies;
- EFC3: Taxes and bureaucracy;
- EFC4: Governmental programs;
- EFC5: Basic school entrepreneurial education and training;
- EFC6: Post school entrepreneurial education and training;
- EFC7: R & D transfer;
- EFC8: Commercial and professional infrastructure;
- EFC9: Internal market dynamics;
- EFC10: Internal market openness;
- EFC11: Physical and services infrastructure;
- EFC12: Cultural and social norms.

EFCs are calculated as the mean of the ordinal perceptions of national experts on entrepreneurship, for each country, on several subjects, ranging between 1 and 9. The 12 EFCs are used as independent variables, and as dependent variables, both EI and TEA are used separately.

There were countries for which information concerning some years was not available. Hence, the data consist of unbalanced repeated measurements of EFCs, EI and TEA over time. For example, for Czech Republic and Serbia, the data were only available in three time points, contrasting with the 18 measures available for Germany, Ireland and Slovenia. Table 1 shows the distribution of the number of time points available for each country.

Table 1. Countries and respective number of years for which data are available.

Country	#	Country	#
Spain	19	Poland, Portugal	10
Croatia, Germany, Ireland, Slovenia	18	Slovakia	9
United Kingdom, Greece	17	Bosnia and Herzegovina, Iceland	8
Finland, Italy, Netherlands, Norway	16	Luxemburg, North Macedonia	7
Switzerland	15	Austria, Estonia	6
Hungary, Sweden	13	Romania	5
Belgium, Russia	12	Cyprus, Bulgaria, Lithuania	4
Denmark, France, Latvia	11	Czech Republic, Serbia	3

2.2. Methodology

This section describes the methods applied to understand the relation between the 12 EFCs and EI, as well as the relation with TEA. A first visual inspection of the evolution through the years of the EFCs, depicted in Figure 1, shows a possible change in the structure

of the EFCs, especially until and after 2008. Hence, before analysing the impact of each EFC on EI and on TEA, a FA was applied to identify whether the correlation structure of the EFC indicators changes throughout time.

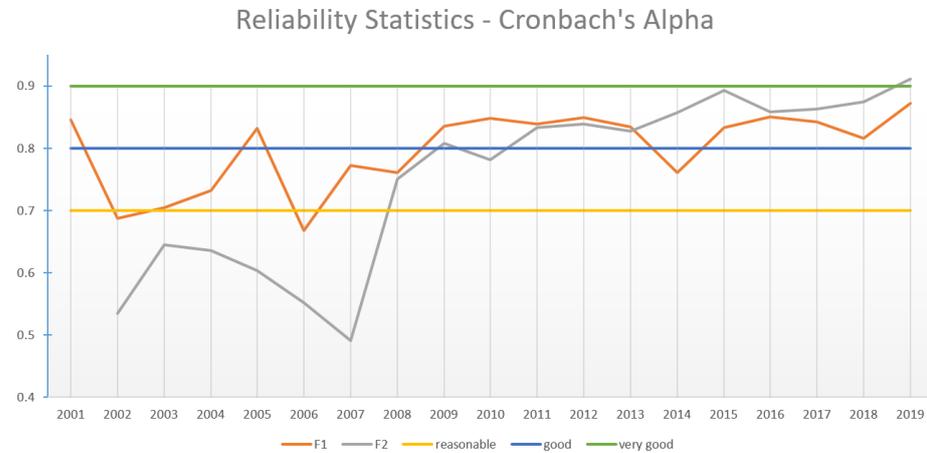


Figure 1. Reliability statistics—Cronbach’s alpha-number of factors—2.

Subsequently, longitudinal regression models were applied for both dependent variables—EI and TEA.

For the FA analysis, the software IBM SPSS v.28 was used. For the dynamic longitudinal analysis, the STATA software was used, in particular the command `xtabond2` created by [17]. Next, both FA and longitudinal analysis will be explained in detail.

2.2.1. Feature Extraction

In the age of big data, an appropriate selection of attributes is essential for several reasons, some of which are next enumerated. The knowledge induced by data analysis algorithms on a smaller number of attributes is often more understandable. Moreover, several algorithms present worse performance when using a large number of variables, and the attributes’ selection can improve it. Thus, when large multivariate data sets are analysed, it is often desirable to reduce its dimensionality. In addition, some domains have a high cost of data collection, in which cases, attribute selection can reduce costs. For a review on the recent developments of feature extraction methods such as PCA and FA, see [18].

In statistics, machine learning and information systems, dimensionality reduction (or dimension reduction) is the process of reducing the number of random variables into consideration in order to obtain a set of main variables. The main approaches for dimension reduction can be divided into: feature selection and extraction. The selection methods try to choose the most representative variables as variables to consider. The extraction methods use the information contained in the data to create another data set with a smaller dimension, while retaining as much information from the original data as possible.

PCA is an extraction technique. It replaces the p original variables, $X_i, i = 1, \dots, p$, by a smaller number, $q < p$, of the derived variables $Y_j, j = 1, \dots, q$ —the principal components (PC). The PC are non-correlated linear combinations of the original variables, i.e.,

$$\begin{aligned}
 Y_1 &= a_{11}X_1 + a_{21}X_2 + \dots + a_{p1}X_p \\
 Y_2 &= a_{12}X_1 + a_{22}X_2 + \dots + a_{p2}X_p \\
 &\vdots \\
 Y_q &= a_{1q}X_1 + a_{2q}X_2 + \dots + a_{pq}X_p.
 \end{aligned}
 \tag{1}$$

PCA aims to represent (or describe) an initial number of variables by a smaller number of hypothetical variables. Thus, it identifies new variables (in smaller numbers) without

a significant lost of information. The components are calculated in descending order of importance, with the first explaining as much of the data variability as possible, i.e.,

$$V[Y_1] \geq V[Y_2] \geq \dots \geq V[Y_q].$$

Often, it is possible to retain most of the variability in the original variables, with q much smaller than p [19].

PCA is based on the analysis of linear correlations between variables, which tries to identify sets of variables that are highly correlated. It allows one to conclude whether it is possible to explain the pattern of correlations through a smaller number of variables. As each linear combination explains as much as possible of the unexplained variance and it has to be orthogonal to any other combination already defined, the set of all combinations found constitutes a unique solution.

Therefore, it is an exploratory analysis technique used to reduce the size of the data, and to identify latent factors.

The representation by PC does not require assumptions on the probability distribution of the variables. However, PCA is largely dependent on the existence of variables with different scales. Therefore, variables must be standardised prior to applying PCA.

PCA implies the absence of error; the latent variables are linear combinations of the initial variables.

In Factor Analysis (FA), each observed variable is described as a function of the common factors retained, and the error is presented, because usually a smaller number of factors is retained when compared with the initial number of variables.

In FA, only the common variation, shared by all the variables, is retained in each factor, while in PCA, the total variation is present in the set of original variables.

Thus, a rotation process is considered in FA. A rotation is a linear transformation that is performed on the component solution for the purpose of making the solution easier to interpret. Interpreting a rotated solution is to identify the concept measured by each of the retained components, also designated as a construct or latent variable.

The number of factors to retain can be defined using several criteria, for example, the factors with an eigenvalue larger than 1 and the percentage of variation of the original variables retained, among others.

PCs, Y_j , are expressed as linear combinations of the original variables, X_i (see Equation (1)). Thus, the error is absent.

In contrast, in FA, each observed variable, X_i , is described as a function of the common factors, FC_k , $k = 1, \dots, m$, with $m < p$:

$$\begin{aligned} X_1 &= \alpha_{11}FC_1 + \alpha_{21}FC_2 + \dots + \alpha_{mp1}FC_{pm} + e_1 \\ X_2 &= \alpha_{12}FC_1 + \alpha_{22}FC_2 + \dots + \alpha_{mp2}FC_{pm} + e_2 \\ &\vdots \\ X_p &= \alpha_{1p}FC_1 + \alpha_{2p}FC_2 + \dots + \alpha_{mp}FC_{pm} + e_p. \end{aligned} \tag{2}$$

However, the factors are easier to interpret and, therefore, facilitate the definition of latent variables in the data.

Data Adequacy

It is not always appropriate to use FA. In fact, exploratory FA is only useful if the matrix of population correlation is statistically different from the identity matrix [20]. It is, then, necessary that the variables are correlated. The Bartlett Sphericity test can be used to test the equality of the correlation matrix to the identity matrix. Only if there is a statistical difference between them is the FA is useful in the estimation of common factors.

The sample adequacy for FA application can be measured by the Kaiser–Meyer–Olkin (KMO) coefficient. This coefficient is based on the partial correlations between the variables and provides information on whether variable selection and sample size are suitable for FA. Sample adequacy can be classified as: incompatible, for $KMO < 0.50$; bad, for KMO

$\in [0.50, 0.60]$; standard, for $KMO \in [0.60, 0.70]$; medium, for $KMO \in [0.70, 0.80]$; good, for $KMO \in [0.80, 0.9]$; and very good, for $KMO > 0.90$ [21].

The Measure of Sample Adequacy (MSA) of each item/variable must be >0.50 . If the variable has a MSA lower than this value, it must be removed from the analysis, as it is not sufficiently correlated with the others and will not be suitable for the analysis.

The communalities, computed in both procedures, PCA and FA, refer to the percentage of the accounted variance of the observed variable retained in the components (or factors). A given variable will display a large communality if it loads heavily on at least one of the retained components.

Retaining Factors

There are several methods that can be used to select the appropriate number of components to retain in FA [22]. The most used are: the Kaiser criterion, which proposes to retain the factors with an eigenvalue larger than 1; the observation of the Scree plot, by evaluating when there is a substantial decline in the magnitude of the eigenvalues; or by specifying a limit value for the cumulative percentage of variance explained by the factors (which are usually larger than 70%). In several research areas, the interpretability criteria are perhaps the most important criterion for determining the number of components.

Reliability

To measure the reliability of a factor, the Cronbach's alpha can be used. This is the most widely used objective measure of reliability. Lee Cronbach established the Cronbach's alpha in 1951 [23] to provide a measure of a test or scale's internal consistency, and it is determined by a number between 0 and 1. The amount to which all elements measure the same concept or construct is referred to as internal consistency. The value of Cronbach's alpha rises as the variables become more linked. A high Cronbach's alpha coefficient, on the other hand, does not automatically imply a high level of internal consistency. More comparable items testing the same notion should be added to raise Cronbach's alpha.

Typically, the acceptable values of Cronbach's alpha range from 0.70 to 0.95 [24].

2.2.2. System GMM

This study analyses repeated measures of EFCs through time for each country, hence longitudinal models (also known as panel models) are applied.

Contrary to cross-sectional methods that aggregate time-series, which may conceal an underlying dynamic (macroeconomic or of other nature), longitudinal models offer the possibility to account for, and investigate, the heterogeneity that may occur between economies.

The standard Fixed Effects Model (FEM) and the Random Effects Model (REM) control for the existence of a bias related to heterogeneity across economies and time. FEM considers individual, unobserved characteristics (i.e., unobserved heterogeneity). The REM model specification assumes that group effects follow a normal distribution over all the groups. However, these models do not overcome the endogeneity problem due to the potential correlation between one or several explanatory variables and the residuals. One strategy to overcome the endogeneity concern could be the Instrumental variable (IV) estimator; however, the main challenge is obtaining valid instruments applicable to panel analyses [25] and theoretically validating them.

In addition to the aforementioned problem, there is another disadvantage in using the FEM and REM models, which is that they are static and do not allow the unbiased estimation of parameters when including past (lagged) values of the dependent variable in the model.

In fact, recently, studies that analyse EI and TEA (e.g., [26–29]), have taken into account the dynamic effect of past values of the variables in their estimations, finding a statistically significant effect. Thus, it was decided to consider in the present analysis the effect of the lagged values of EI and TEA in the estimations.

Because the lagged dependent variable is dependent (correlated) on the error term, a dynamic longitudinal model should be considered. In fact, using the pooled OLS estimation technique yields inconsistent estimates. Furthermore, using the FEM estimator to transform the data to remove the fixed effects does not completely eliminate the inconsistency, because the transformed lagged dependent variable still depends on the error term [30,31].

Moreover, the data under analysis is extremely unbalanced, as there is no information on the variables of interest for all countries in all moments (years). This unbalanced nature of the data requires that the longitudinal dynamics of the expected values of EI and TEA are analysed through a model that overcomes this issue.

Thus, in this analysis, the General Method of Moments (GMM) estimator, proposed by [32,33], is used. This model controls for possible endogeneity and unobserved heterogeneity. This technique uses lags of the dependent variables as explanatory variables to capture the effect of past values of the dependent variable. In this, lagged dependent variable values are used as (internal) instruments to control this endogenous relationship. The GMM model eliminates endogeneity by “internally transforming the data”. This transformation is a statistical process that subtracts a variable’s past value from its present value [17]. As a result, the number of observations is reduced, and this process (internal transformation) improves the GMM model’s efficiency [34].

However, to prevent potential data loss, as we deal with unbalanced data, the second-order transformation (two-step GMM) was considered, as recommended by [32]. This transformation applies “forward orthogonal deviations”, subtracting the average of all future available observations of a particular variable from its current value [17], instead of only subtracting from the previous observation (as in the first-step procedure).

The GMM model considered can be parameterised, as follows:

$$Y_{it} = \varphi Y_{it-1} + \beta X_{it} + (\eta_i + \varepsilon_{it}), \quad (3)$$

where Y_{it} is the dependent variable (in this case, IE or TEA), X_{it} is the vector of independent variables, φ is the autoregressive parameter, η_i is the unobserved fixed effect and ε_{it} is the idiosyncratic shocks normally distributed with zero mean and constant variance. The error term follows the error component structures, in which $E(\eta_i) = 0$; $E(\varepsilon_{it}) = 0$ and $E(\eta_i \varepsilon_{it}) = 0$, $E(\varepsilon_{it} \varepsilon_{its}) = 0$.

Additionally, to capture the overall economic and social context, it is controlled for per-capita GDP (following the works of [35–37]), for size of country’s population, as this is likely to affect the number of people available to work in the labour force, as well as the country’s entrepreneurship rates [38]. Furthermore, year dummy variables (time-specific effects) were included to reduce the influence of cross-sectional error dependence in the dynamic panel model.

System GMM uses moment conditions that are functions of the model parameters and the data, such that their expectation is zero at the parameters’ true values. It controls for endogeneity of the lagged dependent variable in a dynamic panel model—when there is correlation between the explanatory variable and the error term in a model; omitted variable bias; unobserved heterogeneity; and measurement errors.

This method consists of a system of two equations: an original equation, expressed in the levels’ form with the first differences as instruments. It transforms all aggressors through differentiation, thus removing fixed effects (that do not vary over time); and using a transformed equation, expressed in a first-differenced form with levels as instruments. To clarify:

- Equations in differences: $\Delta Y_t = Y_{t-2} - Y_{t-1}$
- Equations in levels: $Y_t = Y_{t-1}; Y_{t-(n-1)} = Y_{t-n}$

where Y_{t-n} is the instrument of $Y_{t-(n-1)}$.

The calculation of the system GMM estimator is based on a stacked system comprising $0.5(T + 1)(T - 2)$ moment conditions, for which instruments are observed (where T represents the total number of years, in this specific case). The Arellano–Bond test for serial

correlation is used to test for second-order serial correlation in the first-differenced residuals for model diagnostics. The null hypothesis states that the residuals are uncorrelated serially. If the null hypothesis cannot be rejected, it demonstrates that there is no second-order serial correlation and that the GMM estimator is reliable. In addition, the Hansen J-test is used to assess the null hypothesis of instrument validity as well as the validity of the additional moment restriction required for system GMM. The instruments are valid if the null hypothesis is not rejected. The variance inflation factor (VIF), an indication of how much the standard error has inflated, was used to calculate collinearity diagnostics for the regression equation estimated (results are not presented due to space restrictions, but available under request).

3. Results

3.1. Feature Extraction—Factor Analysis

The results of applying a FA to the full data set, the EFC in European countries during two decades (2002–2019), considering varimax rotation with Kaiser Normalisation, are presented in Table 2.

The obtained results demonstrate a KMO = 0.898, corresponding to a very good adequacy of the sample for FA application. The MSA of each independent variable are larger than 0.75, indicating that all of them contribute to a significant difference between the correlations’ matrix and the identity matrix. Using the criterion of the eigenvalues larger than 1, two factors are extracted and approximately 64% of the global initial variance of the variables is explained. The communalities of the variables are larger than 0.437.

Table 2. Factor analysis—full data set: 2000–2019.

EFC	Description	MSA	Rotated Component Matrix		Communalities
			F1	F2	
4	Governmental programs	0.914	0.774		0.730
7	R&D transfer	0.919	0.738		0.765
9	Internal market dynamics	0.750	−0.718		0.576
2	Governmental support and policies	0.915	0.700		0.670
8	Commercial and professional infrastructure	0.902	0.698		0.653
10	Internal market openness	0.928	0.650		0.690
1	Financing for entrepreneurs	0.911	0.626		0.495
11	Physical and services infrastructure	0.891	0.608		0.437
12	Cultural and social norms	0.839		0.860	0.779
5	Basic school entrepreneurial education and training	0.893		0.784	0.630
6	Post school entrepreneurial education and training	0.905		0.653	0.514
3	Taxes and bureaucracy	0.905		0.639	0.689
% of variance			35.367	28.206	63.573
Cronbach’s alpha			0.810	0.782	KMO
Number of items			8	4	0.898

The first factor explains approximately 36% of the global variance of the variables and corresponds to the structural conditions of the countries (EFCs 1, 2, 4, 7, 8, 9, 10 and 11) (Table 2). The second explains approximately 28% of the global variance of the variables and it corresponds to norms and education (EFCs 3, 5 and 6), taxes and bureaucracy (EFC 12). The reliability of Cronbach’s alpha values for each factor is acceptable, as they are larger than 0.7.

The EFCs’ indicators, which are combined in the second component, include very different subjects and are poorly interpretable when combined into a single factor. However, this behaviour is different over time. This is presented in Table 3 and discussed in the corresponding text. In Table 4, it can also be observed that in the beginning (first few years) some of these EFCs have MSAs less than 0.4, indicating that they are not very correlated with the others in the analysis. Reliability statistics—Cronbach’s alpha, for two

factors (Table 5) also illustrates reliability values that are larger in the last few years. Thus, effectively, this second factor, although to present acceptable reliability, is not evidence of their consistency in terms of real interpretation and constitution.

Table 3. Factor analysis, per year, for two factors.

EFC	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	#
1	F1	18																	
2		F1		F1	F1	F1	F1	F1		F1	F1	14							
3			F1				F1				F1	F1		F1	F1		F1	F1	8
4			F1		F1	F1	F1		F1	F1	15								
5	F1							F1				F1		F1	F1	F1	F1	F1	8
6	F1	F1							F1		F1		F1	F1	F1	F1	F1	F1	9
7			F1		F1	F1	16												
8			F1		F1		F1	F1	F1	F1	F1	F1	12						
9				F1			F1						9						
10	F1			F1			F1	F1	F1		F1	13							
11			F1	F1	F1		F1	F1	F1	F1	F1				F1	F1	F1	F1	13
12	F1										F1	F1		F1	F1	F1	F1	F1	8

Table 4. Factor analysis for each year from 2002 to 2019.

Year	KMO	Total TVE	# of Factors	TVE		MSA < 0.4	Com. < 0.5
				3 Factors	2 Factors		
2002	0.52	81%	4	72%	58%	5, 6, 11, 12	6
2003	0.56	77%	4	68%	54%	5	–
2004	0.49	71%	3	71%	60%	5, 12, 11	1, 11
2005	0.50	76%	3	76%	64%	6	–
2006	0.71	76%	3	76%	64%		6
2007	0.64	81%	4	76%	65%		6
2008	0.63	76%	3	76%	73%	6	–
2009	0.68	75%	2	81%	75%	6	–
2010	0.57	75%	3	75%	65%	9, 6	–
2011	0.63	80%	3	80%	69%	9	–
2012	0.84	80%	3	80%	71%	–	–
2013	0.86	70%	2	77%	70%	–	–
2014	0.84	76%	3	76%	67%	–	–
2015	0.75	80%	3	80%	70%	9	–
2016	0.80	68%	2	75%	68%	–	–
2017	0.77	74%	2	81%	74%	–	–
2018	0.68	75%	3	75%	65%	9	–
2019	0.82	74%	2	81%	74%	–	–

TVE—Total Variance Explained for number of factors.

Table 5. Reliability statistics—Cronbach’s alpha, for two factors.

Year	F1	F2	# Countries	Year	F1	F2	# Countries
2000	0.748	*	11	2010	0.848	0.782	21
2001	0.846	**	14	2011	0.839	0.833	22
2002	0.687	0.535	16	2012	0.849	0.839	28
2003	0.705	0.645	17	2013	0.834	0.827	28
2004	0.732	0.636	15	2014	0.761	0.858	29
2005	0.832	0.604	18	2015	0.833	0.893	24
2006	0.668	0.552	18	2016	0.851	0.859	25
2007	0.772	0.491	17	2017	0.842	0.863	20
2008	0.761	0.75	14	2018	0.816	0.875	20
2009	0.836	0.808	19	2019	0.872	0.911	22

Missing data: * EFC12; ** EFC6.

In Table 5 (and in Figure 1), the reliability of Cronbach’s alpha results, for applying FA to the European countries’ EFCs, are presented. For each year of the data set, two factors and a varimax rotation with Kaiser Normalization (FA), are considered. The results are between 0.491 and 0.911. These values demonstrate a possible modification in the structure of the EFCs, from year to year, especially until and after 2008. After 2008, all the reliability

Cronbach's alpha values are larger than 0.7, indicating a more stable structure of the EFCs in two factors with good reliability. Note that, for each year, the number of countries with data available is not the same.

The results obtained, for each one year of the data set, considered as criterion to retain the factors that the eigenvalues > 1 , are presented in Table 4. The number of factors defined using this criterion varies between two and four factors. After 2008, the values obtained are more stable, with larger KMO values, and defining two or three factors. The EFCs with $MSA < 0.4$ and with communalities < 0.5 are also fewer after 2018.

Despite reasonable reliability, the structure obtained using the exploratory FA for each year is quite different. However, the main items remain the same. This is illustrated in Table 3, where F1 indicates that the EFC in that row is in Factor 1, and the last column presents the number of years for which the EFC is in Factor 1. For example, EFC1 is in Factor 1 for all years from 2002 to 2019, thus in 18 years. In general, taxes and bureaucracy (EFC3), basic (EFC5) and post (EFC6) school entrepreneurial education and training and cultural and social norms (EFC12), remain in Factor 2. These correspond to norms and education. The remaining EFCs (blue rows) remain in Factor 1, which corresponds to structural conditions of countries.

3.2. System GMM Results

The estimations from the six longitudinal models adjusted are presented in Table 6.

For the time horizon of 2001–2008, EI is negatively associated with the perception regarding the extent to which public policies support entrepreneurship as a relevant economic issue (EFC2), and positively associated with the perception regarding the presence of property rights, commercial, accounting and other legal and assessment services and institutions that support or promote SMEs (EFC8). Moreover, TEA is only positively associated with the perception regarding the extent to which public policies support entrepreneurship (taxes or regulations are either size-neutral or encourage new and SMEs) (EFC3).

For the time horizon of 2009–2019, EI is negatively associated with the perception regarding ease of access to physical resources—communication, utilities, transportation, land or space—at a price that does not discriminate against SMEs—EFC11. TEA is positively associated with the perception regarding the presence and quality of programs directly assisting SMEs at all levels of government (national, regional and municipal) (EFC4); with the perception regarding the extent to which training in creating or managing SMEs is incorporated within the education and training system in higher education such as vocational, college, business schools, etc. (EFC6); and the perception regarding the extent to which social and cultural norms encourage or allow actions leading to new business methods or activities that can potentially increase personal wealth and income (EFC12). Moreover, TEA is negatively associated with the perception regarding the extent to which national research and development will lead to new commercial opportunities and is available to SMEs (EFC7).

For all estimations, the past values of EI and TEA are positively and significantly associated with the present values of EI and TEA (p -value < 0.01), respectively. The results for the Hansen test indicate that, for all specifications, the null hypothesis is not rejected, which means the instruments are valid and the model is specified correctly. Regarding the reported p -values for AR(1) and AR(2) Arellano–Bond autocorrelation tests, as expected, there is first-order serial correlation and no evidence of significant second-order serial correlation. The significance of F statistics also demonstrates that all independent variables jointly and significantly explain the model at a 5% significance level.

Table 6. Results from system GMM estimations.

Variables	Dependent Variable EI			Dependent Variable TEA		
	Full Sample	2001–2008	2009–2019	Full Sample	2001–2008	2009–2019
EI($t - 1$)/ TEA($t - 1$)	0.785 *** (0.038)	0.714 *** (0.049)	0.760 *** (0.053)	0.632 *** (0.061)	0.487 *** (0.140)	0.582 *** (0.085)
EFC1	−1.000 (0.754)	−1.121 (1.038)	−0.973 (0.970)	−0.171 (0.407)	−0.138 (0.695)	0.126 (0.515)
EFC2	0.170 (0.484)	−1.766 ** (0.789)	0.873 (0.606)	−0.315 (0.288)	−1.129 (0.698)	−0.305 (0.329)
EFC3	0.495 (0.481)	0.520 (0.633)	0.350 (0.611)	0.118 (0.292)	1.100 * (0.564)	−0.331 (0.332)
EFC4	0.373 (0.469)	0.244 (0.938)	0.899 (0.791)	0.605 (0.359)	0.763 (0.819)	0.677 * (0.396)
EFC5	−0.206 (0.499)	−1.098 (1.255)	0.379 (0.809)	0.255 (0.294)	−0.071 (0.813)	0.237 (0.423)
EFC6	0.936 (0.625)	0.292 (0.783)	1.114 (1.156)	0.539 (0.343)	0.093 (0.950)	1.035 * (0.526)
EFC7	−0.367 (0.816)	1.108 (2.358)	−0.895 (1.159)	−1.724 *** (0.600)	−0.374 (1.301)	−2.631 *** (0.709)
EFC8	1.337 * (0.665)	3.454 * (1.815)	1.327 (0.827)	0.107 (0.442)	−0.789 (1.121)	0.430 (0.492)
EFC9	0.632 (0.472)	0.965 (0.833)	0.463 (0.586)	−0.209 (0.205)	−0.111 (0.491)	−0.244 (0.254)
EFC10	−0.079 (0.715)	0.202 (1.193)	−0.786 (0.980)	0.448 (0.473)	−0.253 (1.010)	0.744 (0.566)
EFC11	−0.288 (0.323)	0.317 (0.647)	−0.969 * (0.520)	0.233 (0.234)	0.154 (0.485)	−0.019 (0.300)
EFC12	−0.479 (0.549)	0.427 (1.094)	−0.560 (0.763)	0.909 * (0.453)	0.855 (0.878)	1.227 * (0.611)
<i>Control</i>						
GDP_pc	−0.000 *** (0.000)	−0.000 ** (0.000)	−0.000 *** (0.000)	−0.000 (0.000)	−0.000 (0.000)	−0.000 (0.000)
<i>Instrumental</i>						
Population_total	−0.000 *** (0.000)	−0.000 *** (0.000)	−0.000 *** (0.000)	−0.000 *** (0.000)	−0.000 *** (0.000)	−0.000 *** (0.000)
Year dummies	yes	yes	yes	yes	yes	yes
F statistic	533.37 ***	3083.76 ***	533.37 ***	650.91 ***	13624.15 ***	1428.25 ***
Hansen test statistic	0.98	0.98	0.98	3.0	3.0	3.40
<i>Arellano–Bond statistic</i>						
AR(1)	−3.48 ***	−2.94 ***	−3.48 ***	−3.58 ***	−2.90 ***	−3.33 ***
AR(2)	1.34	0.49	1.34	0.36	−0.79	0.77

* p -value < 0.10; ** p -value < 0.05; *** p -value < 0.01.

4. Conclusions and Future Work

To answer the two research questions: “How does the EFCs influence the Entrepreneurship Intentions?” and “How does the EFCs influence the early stage Entrepreneurial Activity?”, a set of statistical methods that allow and take into account the expected dynamic in the time and intra-country of these indicators were used.

A feature extraction, namely FA, applied as a first exploratory analysis, allowed us to understand that two major time horizons should be considered, i.e., 2001–2008 and 2009–2019. Additionally, the fact that the results of the system GMM estimations are different between the full time horizon (2001–2019), the 2001–2008 and the 2009–2019 time horizons reveal the pertinence of the FA.

The results demonstrate that not all EFCs indicators are associated with entrepreneurship intentions and early stage entrepreneurial activity. Furthermore, the associations differ between the two time horizons considered, namely: until 2008 and after.

For the first time horizon, conditions such as the perception regarding the presence of property rights, commercial, accounting and other legal and assessment services and institutions that support or promote SMEs and the extent to which public policies support

entrepreneurship positively affect EI and TEA. The perception regarding the extent to which public policies support entrepreneurship as a relevant economic issue negatively impacts EI, but has no significant impact on TEA.

For the 2009–2019 time horizon, only the perception regarding ease of access to physical resources—communication, utilities, transportation, land or space—at a price that does not discriminate against SMEs, negatively affects EI, but not TEA. TEA is only negatively affected by the perception regarding the extent to which national research and development will lead to new commercial opportunities and is available to SMEs. Furthermore, in this time period, the TEA is positively affected by perception regarding the presence and quality of programs directly assisting SMEs at all levels of government (national, regional and municipal); with the perception regarding the extent to which training in creating or managing SMEs is incorporated within the education and training system in higher education such as vocational, college, business schools, etc.; and the perception regarding the extent to which social and cultural norms encourage or allow actions leading to new business methods or activities that can potentially increase personal wealth and income.

The difference between the impact of EFCs on EI and on TEA enhances the importance of combining the two statistical procedures considered in the present analysis—the FA and the dynamic longitudinal analysis—in studies on this matter.

On all longitudinal dynamic estimations, past values of EI and TEA are significantly associated with present values which, once again, corroborates the dynamic aspect of how structural conditions affect EI and TEA.

The analysis of the relation between EFC and other entrepreneurial behaviour and attitudes, as well as the exploration of further control variables for model robustness and to study the primary data instead of the GEM indicators, are future work possibilities. For example, as entrepreneurship may be related to human ability and propensity to take risks, these qualities could be controlled in future analysis. Additionally, the implementation of a longitudinal Principal Component Analysis would also reinforce the present analysis.

Author Contributions: This work was conducted by the four authors in collaboration through joint and distributed tasks. Joint tasks included conceptualization, organization of the paper, state-of-the-art investigations, definition of the methodology, collection of the resources, interpretation, writing—original draft preparation and writing—review and editing. Major contribution in software implementation, validation and visualization was given by A.B. and A.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work has been supported by national funds through FCT—Fundação para a Ciência e Tecnologia through project UIDB/04728/2020.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data used in this work is available on the GEM project website at <https://www.gemconsortium.org/data>, accessed on 29 June 2020.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Farinha, L.; Lopes, J.; Bagchi-Sen, S.; Sebastião, J.R.; Oliveira, J. Entrepreneurial dynamics and government policies to boost entrepreneurship performance. *Socio-Econ. Plan. Sci.* **2020**, *72*, 100950. [CrossRef]
2. Fernandes, A.J.; Ferreira, J.J. Entrepreneurial ecosystems and networks: A literature review and research agenda. *Rev. Manag. Sci.* **2022**, *16*, 189–247. [CrossRef]
3. Hong, S.; Robert Reed, W.; Tian, B.; Wu, T.; Chen, G. Does FDI promote entrepreneurial activities? A meta-analysis. *World Dev.* **2021**, *142*, 105436. [CrossRef]
4. Thompson, E.R. Individual entrepreneurial intent: Construct clarification and development of an internationally reliable metric. *Entrep. Theory Pract.* **2009**, *33*, 669–694. [CrossRef]
5. Valliere, D. Reconceptualizing entrepreneurial framework conditions. *Int. Entrep. Manag. J.* **2010**, *6*, 97–112. [CrossRef]
6. Orobia, L.A.; Tusiime, I.; Mwesigwa, R.; Ssekiziyivu, B. Entrepreneurial framework conditions and business sustainability among the youth and women entrepreneurs. *Asia Pac. J. Innov. Entrep.* **2020**, *14*, 60–75. [CrossRef]

7. Global Entrepreneurship Research Association. *Global Entrepreneurship Monitor 2016/17 Global Report*; Global Entrepreneurship Research Association: London, UK, 2017.
8. Kelley, D.; Bosma, N.; Amorós, J.E. *Global Entrepreneurship Monitor 2010 Executive Report*; Global Entrepreneurship Research Association: London, UK, 2011.
9. Bergmann, H.; Mueller, S.; Schrettle, T. The use of global entrepreneurship monitor data in academic research: A critical inventory and future potentials. *Int. J. Entrep. Ventur.* **2014**, *6*, 242–276. [[CrossRef](#)]
10. Costa e Silva, E.; Correia, A.; Duarte, F. How Portuguese experts' perceptions on the entrepreneurial framework conditions have changed over the years: A benchmarking analysis. *AIP Conf. Proc.* **2018**, *2040*, 110005.
11. Sampaio, C.; Correia, A.; Braga, V.; Braga, A. Environmental variables and entrepreneurship—A study with GEM NES global individual level data. *Int. J. Knowl. Based Dev.* **2016**, *2040*, 4476.
12. Costa e Silva, E.; Correia, A.; Borges, A. Unveiling the dynamics of the European entrepreneurial framework conditions over the last two decades: A cluster analysis. *Axioms* **2021**, *10*, 149. [[CrossRef](#)]
13. Braga, V.; Queirós, M.; Correia, A.; Braga, A. High-growth business creation and management: A multivariate quantitative approach using GEM data. *J. Knowl. Econ.* **2018**, *9*, 424–445. [[CrossRef](#)]
14. Pilar, M.D.F.; Marques, M.; Correia, A. New and growing firms' entrepreneurs' perceptions and their discriminant power in EDL countries. *Glob. Bus. Econ. Rev.* **2019**, *21*, 474–499. [[CrossRef](#)]
15. Levie, J.; Autio, E. Entrepreneurial framework conditions and national-level entrepreneurial activity: Seven-year panel study. In *Proceedings of the Third Global Entrepreneurship Research Conference, San Francisco, CA, USA, 28–31 March 2007*; pp. 1–39.
16. Van Stel, A.; Carree, M.; Thurik, R. The effect of entrepreneurial activity on national economic growth. *Small Bus. Econ.* **2005**, *24*, 311–321. [[CrossRef](#)]
17. Roodman, D. How to do xtabond2: An introduction to difference and system GMM in Stata. *Stata J.* **2009**, *9*, 86–136. [[CrossRef](#)]
18. Jolliffe, I.T.; Cadima, J. Principal component analysis: A review and recent developments. *Philos. Trans. R. Soc. A: Math. Phys. Eng. Sci.* **2016**, *374*, 20150202. [[CrossRef](#)]
19. Jolliffe, I. Principal component analysis. In *Encyclopedia of Statistics in Behavioral Science*; John Wiley & Sons, Ltd.: Hoboken, NJ, USA, 2005.
20. Sarmiento, R.; Costa, V. *Comparative Approaches to Using R and Python for Statistical Data Analysis*; IGI Global: Hershey, PA, USA, 2017.
21. Cureton, E.E.; D'Agostino, R.B. *Factor Analysis: An Applied Approach*; Psychology Press: East Sussex, UK, 2013.
22. Cangelosi, R.; Goriely, A. Component retention in principal component analysis with application to cDNA microarray data. *Biol. Direct* **2007**, *2*, 1–21. [[CrossRef](#)]
23. Cronbach, L.J. Coefficient alpha and the internal structure of tests. *Psychometrika* **1951**, *16*, 297–334. [[CrossRef](#)]
24. Tavakol, M.; Dennick, R. Making sense of Cronbach's alpha. *Int. J. Med. Educ.* **2011**, *2*, 53. [[CrossRef](#)]
25. Bound, J.; Jaeger, D.A.; Baker, R.M. Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *J. Am. Stat. Assoc.* **1995**, *90*, 443–450. [[CrossRef](#)]
26. Mohammadi Khyareh, M.; Rostami, N. The Impact of Macroeconomic Variables on Entrepreneurship: Generalized Method of Moments (GMM). *J. Macroecon.* **2021**, *15*, 30–55.
27. Jafari-Sadeghi, V.; Sukumar, A.; Pagán-Castaño, E.; Dana, L.P. What drives women towards domestic vs international business venturing? An empirical analysis in emerging markets. *J. Bus. Res.* **2021**, *134*, 647–660. [[CrossRef](#)]
28. Afawubo, K.; Noglo, Y.A. ICT and entrepreneurship: A comparative analysis of developing, emerging and developed countries. *Technol. Forecast. Soc. Change* **2022**, *175*, 121312. [[CrossRef](#)]
29. Dutta, N.; Sobel, R.S. Entrepreneurship, fear of failure, and economic policy. *Eur. J. Political Econ.* **2021**, *66*, 101954. [[CrossRef](#)]
30. Nickell, S. Biases in dynamic models with fixed effects. *Econom. J. Econom. Soc.* **1981**, *6*, 1417–1426. [[CrossRef](#)]
31. Anderson, T.W.; Hsiao, C. Formulation and estimation of dynamic models using panel data. *J. Econom.* **1982**, *18*, 47–82. [[CrossRef](#)]
32. Arellano, M.; Bover, O. Another look at the instrumental variable estimation of error-components models. *J. Econom.* **1995**, *68*, 29–51. [[CrossRef](#)]
33. Blundell, R.; Bond, S. Initial conditions and moment restrictions in dynamic panel data models. *J. Econom.* **1998**, *87*, 115–143. [[CrossRef](#)]
34. Wooldridge, J.M. *Introductory Econometrics: A Modern Approach*; Cengage Learning: Boston, MA, USA, 2015.
35. Albulescu, C.T.; Tămășilă, M. Exploring the role of FDI in enhancing the entrepreneurial activity in Europe: A panel data analysis. *Int. Entrep. Manag. J.* **2016**, *12*, 629–657. [[CrossRef](#)]
36. Acs, Z.J.; Desai, S.; Hessels, J. Entrepreneurship, economic development and institutions. *Small Bus. Econ.* **2008**, *31*, 219–234. [[CrossRef](#)]
37. Van Stel, A.; Storey, D.J.; Thurik, A.R. The effect of business regulations on nascent and young business entrepreneurship. *Small Bus. Econ.* **2007**, *28*, 171–186. [[CrossRef](#)]
38. Hechavarría, D.M.; Ingram, A.E. Entrepreneurial ecosystem conditions and gendered national-level entrepreneurial activity: A 14-year panel study of GEM. *Small Bus. Econ.* **2019**, *53*, 431–458. [[CrossRef](#)]