



An inquiry into the drivers of an entrepreneurial economy: A Bayesian clustering approach

Maximo Camacho¹ · Emilio Congregado² · Ana Rodriguez-Santiago²

Accepted: 26 April 2024
© The Author(s) 2024

Abstract

Understanding the worldwide drivers of qualified entrepreneurship is a key issue in economic policy design. To help policy decisions exert their intended impact, we aim to cluster a wide range of countries on the basis of their levels and trends in self-employment productivity using a finite mixture model applied to a new large dataset of 121 countries covering the period of 1991–2019. Our results point to three groups of high-, medium-, and low-productive means and tendencies, the geographical distribution of which suggests that they can be reinterpreted using the three stages of economic development, namely, innovation-, efficiency-, and factor-driven economies. Notably, we find that widespread digitalization and low unemployment enhance the probability of transitioning into a highly productive cluster. However, we failed to find that industry weight or employment protection legislation strictness serve as determinants in the transition between groups. Suggestive rationales for these results and implications for the entrepreneurship policy agenda are also provided.

Keywords Entrepreneurship · Productive self-employment · Model-based clustering · Finite mixture models · Cross-country analysis · Transition probabilities

JEL classification M13 · J24 · C11 · O43

1 Introduction

Policymakers tend to encourage entrepreneurial activity because it is viewed as a key driver of economic growth, job creation, and innovation. Consequently, they implement portfolios of policies to promote entrepreneurship/self-employment and

✉ Ana Rodriguez-Santiago
ana.rodriguez@dege.uhu.es

¹ University of Murcia, Murcia, Spain

² Department of Economics and CCTH, Faculty of Business Administration, University of Huelva, Huelva, Spain

to support small and medium firms as a solution to weak economic performance and deficient job creation. However, as the seminal work of Blanchflower (2004) pointed out, the level of self-employment itself does not guarantee economic growth. In fact, as Poschke (2013) noted, both developed and developing countries sometimes show the same self-employment rates despite having different growth patterns.

Among others, Shane (2009) and Congregado et al. (2010) warn that encouraging more people to become entrepreneurs does not necessarily lead to economic development. The strong negative cross-country association between self-employment and the level of income per capita in both less-developed and developing countries and the mixed evidence regarding the impact of entrepreneurship on growth at the macro level constitute indications of something being wrong in the usual linkage between the size of the aggregate self-employment sector and economic growth, as the works of Pietrobelli et al. (2004), Wennekers et al. (2010), Arin et al. (2015), and Rodríguez-Santiago (2022) have found.

Not trying to be exhaustive, Maloney (2004), Acs (2006), Poschke (2018, 2019), and Allub and Erosa (2019) pointed out that self-employment exhibits substantial heterogeneity and that cross-national differences could be behind this apparent puzzle and suggested examining the relationships between qualified self-employment (rather than the aggregate rate) and economic development. In this context, Stam (2015) and Stam and Van de Ven (2021) investigated the determinants of optimally productive entrepreneurship and the pillars of the entrepreneurial ecosystem, which are particularly important for devising an effective national competitiveness strategy.

This body of literature encompasses three main categories of research focusing on the determinants of self-employment rates at the macroeconomic level. Firstly, studies such as those by Acs et al. (1994) delve into the influence of macroeconomic factors like capital per worker and industry composition. Secondly, research by Pietrobelli et al. (2004), Arin et al. (2015), and Rodríguez-Santiago (2022) examines the impact of income per worker on self-employment rates, highlighting the adverse effects of macroeconomic instability on entrepreneurial activities. Thirdly, investigations by Blanchflower (2000), Centeno (2000), Robson (2003), and Torrini (2005) scrutinize labor market dynamics and regulations, including employment protection legislation, while others such as Fölster (2002), Anokhin and Schulze (2009), Djankov et al. (2010), Estrin et al. (2012), Belitski et al. (2016), and Dutta and Sobel (2016) focus on the role of corruption and taxation. Lastly, various articles, including those by Sobel (2008), Acs et al. (2008), Estrin et al. (2012), Bjørnskov and Foss (2016), and Urbano et al. (2020), explore the influence of institutions and institutional quality on entrepreneurship.

For these reasons, our paper follows this literature by focusing on the analysis of self-employment productivity (output per self-employed worker) with a twofold purpose. Our first purpose is to cluster countries worldwide to identify groups with some degree of similarity regarding their level and trend in self-employment productivity. With this classification, policymakers could examine cluster membership to determine whether their countries have performed on par with other countries in similar economic circumstances and to provide warning of unfulfilled expectations.

Our second purpose is to determine what drives country memberships and, as a result, what characteristics are shared by countries that makes them similar to their

cluster and different from other clusters in terms of self-employment productivity. This analysis may provide implications for country policymakers regarding which policy variables have the greatest degree of influence on country-level self-employment performance and to consider which strategies would be appropriate to promote movements towards more productive clusters.

In our study, we use finite mixture models to analyze the varied landscape of self-employment productivity across different countries. These models offer numerous advantages over alternative approaches, particularly in their capacity to provide inference on individual classifications and overall clustering. The following paragraphs lay out a robust foundation for comprehending the chosen methodology.

To capture the worldwide heterogeneity in self-employment productivity, we rely on finite mixture models. These model-based classification methods exhibit several advantages over their alternatives, which classify according to similarities. First, the clustering in finite mixture models works on a statistical basis, which facilitates the conduction of inference on the estimates for individual classifications and for the clustering as a whole. Second, in the context of finite mixture models, a number of statistical criteria have been developed to objectively assess the optimal number of clusters. Third, exogenous explanatory variables explaining cluster formation can be easily and explicitly incorporated in the clustering procedure. Thus, model-based classification facilitates designing guidelines for policy recommendations based on an analysis of outcomes.

In finite mixture models, each cluster is assumed to have its own density. In our approach, we assume that this density is determined by both the group-dependent level of self-employment productivity and its group-dependent trend, or long-term direction, over the sample period. Thus, the model characterizes the homogeneity within each cluster not only by level of self-employment, which could be viewed as either high, moderate, or low, but also by the intensity in the overall evolution of the data path.

In addition, the model defines the probability that a country belongs to a given group. In this case, we consider a logit-type structure that depends on the influence of unit-specific exogenous variables on cluster membership. According to the literature, we postulate that these exogenous variables characterizing the ex-ante likelihood of membership in a given cluster are of four types. First, following Acs (2006), who consider that improvements in information technologies such as telecommunications may increase the returns to entrepreneurship, we use the World Bank's Digital Adoption Index (DAI).

Second, following Fairlie and Fossen (2020) and Cowling and Wooden (2021), we also consider the country-specific labor market situation as measured by the unemployment rate to be an influencing factor on whether a country is affiliated with a certain cluster of self-employment productivity. Third, in line with Blanchflower and Shadforth (2007), who examine whether self-employment was stimulated in the United Kingdom through changes in the industrial sector, we include the relative weight of the industrial sector. Finally, we follow Centeno (2000) and Robson (2003), who examined the interaction of self-employment and labor market rigidity, to propose the Labor Market Rigidity Index (LAMRIG) as an additional exogenous determinant of cluster membership.

In the empirical analysis, we compile a new large internationally comparable database of 121 countries covering the period from 1991 to 2019 and use the finite mixture model to obtain the following results. First, our data-driven approach points to three distinct groups of countries. The first group characterizes the countries with the highest productivity level and the steepest productivity trend. The second group comprises countries with a medium level of productivity and a flatter trend than that in the first group. The third group characterizes the countries with the lowest productivity level and the least pronounced trend in self-employment productivity.

The main results and contributions of the paper can be summarized in the following statements: firstly, referring to the geographical distribution of groups; secondly, regarding the disparities in self-employment productivity across these groups; and thirdly, concerning the pivotal elements influencing transitions between groups.

In accordance with the resulting geographical distribution of groups, we followed the categories of countries suggested by Porter et al. (2002) to categorize the groups according to their levels of competitiveness across the stages of economic development. The first cluster tends to include most innovation-driven economies with higher wages, level of innovation and associated standards of living. The second cluster includes countries in the efficiency-driven stage, who require the development of more efficient production processes and the ability to harness the benefits of existing technologies. Countries in the factor-driven stage predominate in the third cluster, which is composed of the least developed countries, where subsistence agriculture, extraction businesses, and unskilled labor are prevalent.

Second, despite the significant differences in the levels of self-employment productivity across these three groups, our results do not suggest that they will eventually converge. The reason is that the productivity of countries in the lower productive groups tends to grow over the sample period at a slower rate than that in the higher productive groups. Therefore, the trajectory of the least productive countries will not tend to catch up unless policymakers take measures to close the productivity gap. Thus, doing nothing is the best guarantee of failure in promoting the convergence process.

Third, our research identifies two key elements in the national entrepreneurial ecosystem that can enable less productive countries to reverse this tendency by determining the key factors that influence group membership. In line with the intersections of digital technologies and entrepreneurship that have been documented by Jafari-Sadeghi et al. (2021), our results show that designing a nuanced digital strategy with policies tailored to promote adoption and diffusion of digital technologies is especially important in facilitating the transition to the innovation-driven group.

In addition, we find that unemployment is a barrier to moving countries from the efficiency-driven to the innovation-driven group. In line with Thurik et al. (2008), who detected a dynamic interrelationship between self-employment and unemployment rates, we find that the labor market dynamic is also related to self-employment productivity. We postulate that the structural unemployment rate tends to favor the entry of marginal entrepreneurs who erode average productivity into self-employment. Thus, active labor market policies that are oriented to stimulate the search for salaried work offers hinder the promotion of self-employment among the less productive unemployed and thus appear to be advisable as a strategy for catching-up.

Interestingly, we failed to find that industrialization intensity or the rigidity of employment protection legislation were key elements in transitioning to more productive groups. The former is in line with the findings of Acs and Naudé (2013), who do not see industrial policies as merely functional policies without consideration of firm or entrepreneurial specifics. The latter agrees with Robson (2003), Torrini (2005), and Kanninen and Vesala (2005), who found that employment protection legislation restrictiveness had little impact on aggregate self-employment.

The outline of the rest of the paper is as follows. Section 2 presents the model and provides an overview of the key statistical elements used to understand the empirical results. Section 3 describes the database and presents and discusses the results. Section 4 concludes and provides policy implications and further avenues of future research.

2 Data and methodology

2.1 Data description

In this paper, we focus on the productivity of self-employment as a proxy for the quality of entrepreneurship. In particular, we consider GDP per self-employed person, i.e., the output per self-employed worker, as our measure of productivity. To compare productivity levels across countries, GDP is converted to international dollars using purchasing power parity rates, which account for the differences in relative prices among countries.¹

Self-employed workers are those workers who, working on their own account, with one or a few partners or in a cooperative, hold the type of jobs defined as self-employment jobs. These data are taken from the International Labor Office (ILOSTAT) database. Self-employed workers include the following four subcategories: employers, own-account workers, members of producers' cooperatives, and contributing family workers.

The dataset of covariates that are used in the logistic prior to classifying each country in a specific group is created with four variables meant to capture group-specific differences. The first structural variable captures the labor market situation by using the average unemployment rate provided by the ILOSTAT database. This is measured as the percentage of the total labor force that is without work but have been seeking work in a recent past period and is currently available to work.

The second structural variable, which reflects the level of industrialization, is the average of industry added value as a percentage of GDP, including the ISIC divisions 05-43. These data are taken from World Bank national accounts and OECD National Accounts Statistics. The third structural variable is meant to capture the

¹ In particular, we converted GDP to international dollars using purchasing power parity rates, in constant 2017 international dollars, from the World Development Indicators database, World Bank and Eurostat-OECD PPP Programme. For Canada, we use real GDP at constant national prices in millions 2017 US dollars, which is taken from Penn World Table 10.0.

rigidity of employment protection legislation. For this purpose, we use the average of the market legislation rigidity index, as detailed by Campos and Nugent (2018).

Finally, we use a fourth structural variable to measure the level of digitalization. In particular, we use the Digital Adoption Index provided by the World Bank, which is a composite index measuring the spread of digital technologies in a country across three dimensions of the economy, namely, those of people, government, and business. To facilitate interpretation, the data have been normalized so that countries with values over 0 will be above the sample average, and vice versa.

Estimation in finite mixture models requires handling balanced panels. Therefore, our effective dataset is composed of annual self-employment and the four country-level covariates for a large set of 121 countries, spanning from 1991 to 2019. The list of countries, their code, average GDP by self-employed for the period, and the values of covariates can be found in the Appendix Table 4.

2.2 Model specification

In this paper, we investigate pooling within a time series panel using a finite mixture of an unspecified number of separate distributions.² For this purpose, let $y = \{y_{it}\}$, $i = 1, \dots, N; t = 1, \dots, T$ be a panel, where i and t refer to country-specific self-employment productivity and year, respectively. In addition, we assume that the time series arise from K hidden groups in such a way that all the time series within a certain group are characterized by the same econometric model and depend on the same set of parameters, which are heterogeneous across groups.

The approach used is based on formulating a time series model for each univariate time series $y_i = \{y_{i1}, \dots, y_{iT}\}$ in terms of the group-specific sampling density. For group k , the density is $p(y_i|\vartheta_k)$, and the unknown group-specific parameters $\vartheta = \{\vartheta_1, \dots, \vartheta_K\}$ take values in a parameter space θ . In this case, the same model is valid for all the time series within a given group, although with different parameters across groups. Furthermore, we also assume that the time series are independent within each cluster.

In this context, it is convenient to introduce a latent group indicator S_i , which takes a value out of the discrete set $\{1, \dots, K\}$, indicating to which group the time series belongs; that is, $S_i = k$ indicates that y_i belongs to group k . We assume that $S = (S_1, \dots, S_N)$ are a priori independent. Thus, knowing S_i is equivalent to knowing the group-specific parameters and the density $p(y_i|\vartheta_{S_i})$.

The joint sampling distribution reads as

$$p(y|S, \vartheta) = \prod_{k=1}^K \prod_{i:S_i=k} p(y_i|\vartheta_k). \quad (1)$$

However, an important issue in this specification is that neither the number of clusters nor the group membership are known a priori. In contrast, we use model-based clustering techniques based on Bayesian classification rules to determine K and to estimate the group indicator S_i along with the group-specific parameters $\vartheta_1, \dots, \vartheta_K$ from the data.

² For further details on finite mixture models, see Frühwirth-Schnatter (2006).

To overcome the issue that group membership is unknown in practice, we assume that each time series of self-employment productivity is taken to be a realization of the mixture probability density function of K separate distributions

$$y_{it} \sim \sum_{S_i=1}^K p(y_i | \vartheta_{S_i=k}) \Pr(S_i = k | Z_i, \gamma), \tag{2}$$

where the mixing proportion $\Pr(S_i = k | Z_i, \gamma)$ is the probability that y_i belongs to group k . Thus, the probabilities of group membership are posited to rely on vectors of country-specific variables, Z_i , each comprising g variables, and on the parameter set $\gamma = (\gamma_1, \dots, \gamma_K)$, where each γ_k represents a vector of g elements, with $k=1, \dots, K$. For clustering purposes, each component in mixture Model (2) corresponds to a cluster.

To complete the model specification, given the time series of self-employment productivity of country i , y_i , that belongs to a certain group k , $S_i = k$, we consider that the expected value of each time series of this group is fully characterized by a group-dependent mean and a group-dependent trend. Thus, we model the time-series dynamics of self-employment productivity as

$$y_{it} = \mu^{S_i} + \alpha^{S_i}t + \varepsilon_{it} \tag{3}$$

where the error term is conditionally heteroscedastic, $\varepsilon_{it} \sim N\left(0, \frac{\sigma^2}{\lambda_i}\right)$, and $S_i = 1, \dots, K$.³ For each group, we consider $\vartheta_k = (\mu^k, \alpha^k, \sigma^2)$, and for each country, we consider that the series-specific variance weights, λ_i , where $\lambda = (\lambda_1, \dots, \lambda_N)$ collects all of the weights. For a given cluster i belonging to cluster $S_i = k$, parameter μ^{S_i} represents the base values of self-employment productivity that characterizes the cluster, while parameter α^{S_i} provides the intensity of the increasing or decreasing behavior in the series belonging to the cluster over time.

In addition, we follow Frühwirth-Schnatter and Kaufmann (2008), Kaufmann (2010), and Hamilton and Owyang (2012), and consider a multinomial logit model to include prior information on a particular series in the estimation of the group probability:

$$\Pr(S_i = k | Z_i, \gamma) = \frac{\exp(Z_i' \gamma_k)}{1 + \sum_{l=2}^K \exp(Z_i' \gamma_l)}, \tag{4}$$

where the first group is the baseline group, and we set $\gamma_1 = 0$. We assume that $\gamma = (0, \gamma_2, \dots, \gamma_K)$ are independent of the other parameters of the model.

The vector Z_i , for $i = 1, \dots, N$, includes the g country-specific average features of the labor market structure that determine the classification of the self-employment productivity of country i into a specific group, with $Z' = (Z'_1, \dots, Z'_N)$. The parameters $(\gamma_2, \dots, \gamma_K)$ are unknown but group-specific values, and they allow us to estimate the prior classification probabilities of country i belonging to a group depending on the structural variables Z_i .

³ The advantages of Gaussian mixture modeling are that the estimation methods are well established, and the component distributions are thoroughly understood and thus interpretation of the results is facilitated.

These parameters have a nice interpretation because they determine the intensity of each structural variable for classifying a country into a certain group. If the j -th component of γ_k for country i is positive, then there is an important role for the j -th structural variable of country i in making this country more likely to belong to group k rather than to part of the baseline group. In contrast, if the component is negative, increasing the j -th structural variable of country i increases the probability of this country being reclassified toward the baseline group.

2.3 Model estimation

The model estimation is carried out within a Bayesian framework with the aid of Markov Chain Monte Carlo (MCMC) simulation and data augmentation methods for finite mixture models. Thus, using the information given in the data, the key issue is obtaining a posterior inference on the group indicator, S , the model parameters, ϑ , the series-specific variance weights, λ , and the intensity of the structural variables, γ .

Let us start by assuming that the number of clusters K is known, although we will set a procedure for determining the number of clusters below.

Priors The parameter vector is further broken down into parameter blocks, for all of which we assume standard prior distributions as follows: The prior distribution of the group-specific parameters $(\mu^k, \alpha^k) \sim N(m_0, M_0)$, for $k = 1, \dots, K$; the variance of the error terms and the series-specific variance weights follow inverse gamma and gamma distributions, respectively: $\sigma^2 \sim IG(g_0, G_0)$ and $\lambda_i \sim G\left(\frac{v}{2}, \frac{v}{2}\right)$ for $i = 1, \dots, N$; and the parameters governing the prior group probabilities under the logit structure follow a normal distribution, $\gamma_k \sim N(0, \tau I_g)$, for $k = 1, \dots, K$, where g is the dimension of vectors Z_i .

Estimation The sampling scheme to draw from the posterior follows Frühwirth-Schnatter and Kaufmann (2008) and involves the iteration between the following three steps:

- (i) Classification for fixed parameters. Each time series y_i , with $i = 1, \dots, N$, is classified into one of the K groups by sampling the group indicator S_i from the posterior distribution $Pr(S_i = k | y_i, Z_i, \vartheta, \lambda, \gamma)$, using the sampling density as well as the prior classification probabilities,

$$Pr(S_i = k | y_i, Z_i, \vartheta, \lambda, \gamma) \propto p(y_i | \vartheta_k, \lambda_i) Pr(S_i = k | Z_i, \gamma), \quad (5)$$

for $k = 1, \dots, K$.

- (ii) Estimation for a fixed classification and λ . Conditional on knowing the values of S and λ , sampling $\vartheta_1, \dots, \vartheta_K$ is carried out by sampling the group-specific parameters from the posterior $p(\vartheta_1, \dots, \vartheta_K | S, y, \lambda)$, where each group parameter ϑ_k is estimated by pooling each time series that currently belongs to group k . To sample γ , we follow Scott (2011) and use a Metropolis-Hasting algorithm.
- (iii) Estimation of λ for a fixed S , ϑ and γ . For each $i = 1, \dots, N$, the scale factors $\lambda = (\lambda_1, \dots, \lambda_N)$ are sampled independently from the gamma distributions.

The MCMC estimation procedure described above is repeated M times, and the stacked values of the outcomes of each iteration draw can be used to perform an inference. However, the sampler could present label switching problems, and the finite mixture model must be identified through some inequality constraint on the group-specific parameters. To handle label switching in mixture models, we use the identifiability constraint $\mu_1 > \mu_2 > \dots > \mu_K$ for all $k = 1, \dots, K$. In other words, the constraint implies that the groups are identified by their level of self-employment productivity.

Once the model has been identified, we can perform inference regarding which time series belong to which group by using the posterior classification probability. In particular, we can estimate the posterior probability that a time series y_i belongs to group k from the MCMC draws by averaging over the M iterations,

$$Pr(S_i = k | y_i, Z_i, \vartheta, \lambda, \gamma) \approx \frac{1}{M} \sum_{m=1}^M I_{\{S_i^{(m)}=k\}} \tag{6}$$

2.4 The number of clusters

For exposition purposes, the number of components of the mixture, K , was known. In practice, however, the number of groups will be unknown. To choose the number of groups in a straightforward form, one could select the number of components that maximizes the marginal likelihood from the set $\{1, \dots, K^*\}$, where K^* is an upper bound. However, this method will result in a model with an arbitrarily large number of groups.

For this reason, we consider selecting the model with the number of groups necessary to maximize the quality of the classification by introducing the entropy of the model. If we call EN_j the entropy of a model with a fixed number of j groups, the method would entail selecting K as the model that minimizes

$$EN_j = - \sum_{i=1}^N \sum_{k=1}^j Pr(S_i = k | y_i, Z_i, \vartheta, \lambda, \gamma) \log Pr(S_i = k | y_i, Z_i, \vartheta, \lambda, \gamma), \tag{7}$$

for $j = 1, \dots, K^*$. In this expression, larger entropy values indicate worse clustering solutions in terms of a quality classification, where the value would be 0 for perfect classification.

3 Empirical results

3.1 Model estimation

Estimation is based on the following priors. For group-specific parameters, we use $(\mu_k, \alpha_k) \sim N(0, 1000)$. The priors of the variances are $\sigma^2 \sim IG(1, 1)$, and they are $\lambda_i \sim G(4, 4)$ for the scale parameters. For the parameters of the logistic model, we use the prior $\gamma \sim N(0, \tau I_g)$, with $\tau = 20$ and $g = 4$. For each run of the MCMC sampler, after conducting a burn-in phase of 2000 iterations to remove dependence on the starting condition, 8000 draws are kept to evaluate the estimation.

To select the number of groups, we set $K^* = 6$. Table 1 presents the results of the marginal likelihood and entropy for models with up to six groups. As expected, the likelihood increases with the number of groups. However, the model specification that divides the data into three separate groups is preferred because this model reaches the lowest entropy value (0.38) among model specifications.

Table 2 gives the posterior means of the mean and slope coefficients associated with each of the three idiosyncratic groups, displaying their standard deviations in parentheses indicating that they are all significant at the 5% level. The table shows a division of countries into three distinct groups according to the average and trend of their respective self-employment productivity. The groups are ordered in decreasing order of entrepreneurship productivity levels and trends, with Group 1 being associated with countries with the highest levels of productivity and the steepest tendencies. In Group 2, we find countries with medium levels of productivity and mid-level

Table 1 Marginal likelihood of various model specifications

k	Log-likelihood	Entropy
2	-19654.31	0.41
3	-18480.34	0.38
4	-18073.71	0.96
5	-17301.56	0.45
6	-16875.48	1.88

For a model with k groups, the table shows the marginal log-likelihood and the rate of disorder (entropy)

Table 2 Group-specific model parameters

	S_i		
	1	2	3
μ	318.32 (3.86)	99.32 (2.74)	7.45 (0.53)
α	9.88 (0.24)	2.86 (0.17)	0.27 (0.03)

For groups $S_i = 1, 2, 3$, the table shows the posterior means of estimated group-specific model parameters. Standard errors are in parentheses. Parameters μ represent the base values of self-employment productivity and parameters α represent the trend

tendencies. Finally, Group 3 contains the countries with the lowest productivity levels and the flattest tendencies.

To complete the description of the groups, Table 3 illustrates the individual characteristics that drive group formation. For this purpose, the table shows the posterior means of the estimated logistic coefficients that influence the group probabilities and their standard deviations (in parenthesis), with bold indicating that the coefficients are significant at 95% confidence. The structural variables, which appear in columns, represent the averages of unemployment rate, the value added by industry, the Labor Market Rigidity Index, and the Digital Adoption Index.

In accordance with the parameter estimates, we find an important role of the level of ability of individuals in a country to access and use new information and communication technologies (ICTs) to make the country less likely to be part of Groups 2 and 3. Thus, we identify Group 1 as the group of countries with the highest levels of adoption of digital technologies, which lead to high levels and an upward tendency of productive self-employment. To examine this finding more deeply, Fig. 1a shows the prior probability that country i belongs to Group 1, which is conditional on the structural variables, $Pr(S_i = 1|Z_i, \gamma)$, as a function of the Digital Adoption Index. The figure reveals the positive relationship between the adoption of digital technologies and the probability of being classified in the group with the highest level and steepest tendency of self-employment productivity.

In addition, Table 3 shows that countries with high unemployment rates have lower odds of being classified in the first group than of belonging to the second group. To illustrate this finding, Figure 1 (b) shows that the prior probability that country i belongs to Group 2, which is conditional on the structural variables, $Pr(S_i = 2|Z_i, \gamma)$, is positively correlated with unemployment.

The figures in Table 3 also point to a very interesting finding. The parameters governing the membership probabilities that relate to industry size and the rigidity of employment protection legislation are not statistically significant. Thus, neither the value added by industry nor that added by labor market rigidity seem to play a statistically significant role in group formation.

To interpret the dynamics of group membership, Fig. 2 sketches the geographical distribution of the three groups based on the country's highest posterior probabilities. In particular, country i is classified in Group k if

Table 3 Parameters governing the membership probabilities

S_i	Unemployment	Industry	LAMRIG	DAI
2	0.85 (0.37)	0.34 (0.29)	0.21 (0.29)	-0.94 (0.35)
3	-0.49 (0.50)	0.22 (0.36)	0.34 (0.40)	-4.72 (0.79)

For groups $S_i = 2, 3$, the table shows the posterior means of estimated logistic coefficients. Standard errors are in parentheses, with bold indicating statistical significance at 95% confidence. Industry is value added by industry as a percentage of GDP; *LAMRIG* Labor Market Rigidity Index, *DAI* Digital Adoption Index

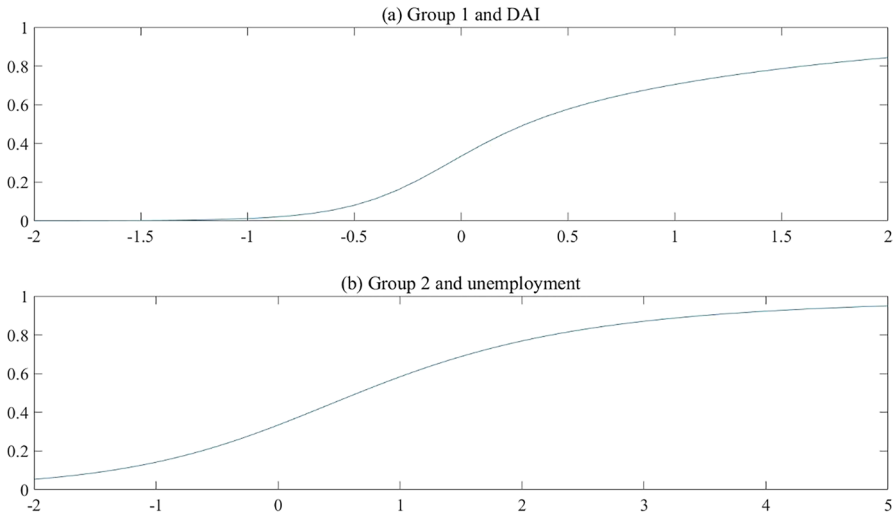


Fig. 1 Membership probabilities. Panel (a) shows the prior probability that country i belongs to Group 1, which is conditional on the structural variables Z_i , $Pr(S_i = 1|Z_i, \gamma)$, as a function of the Digital Adoption Index (DAI). Panel (b) shows the prior probability that country i belongs to Group 2, which is conditional on the structural variables Z_i , $Pr(S_i = 2|Z_i, \gamma)$, as a function of unemployment rate

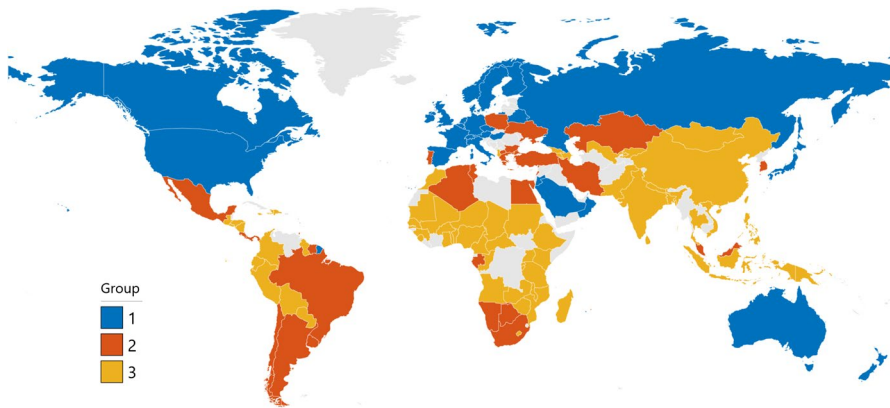


Fig. 2 Group membership. The map shows the geographical distribution of group membership, where country i is classified in Group k if $Pr(S_i = k|y_i, Z_i, \vartheta, \lambda, \gamma) > Pr(S_i = j|y_i, Z_i, \vartheta, \lambda, \gamma)$, with $j \neq k$

$Pr(S_i = k|y_i, Z_i, \vartheta, \lambda, \gamma) > Pr(S_i = j|y_i, Z_i, \vartheta, \lambda, \gamma)$, with $j \neq k$. A visual examination of the map allows us to identify the highly productive countries of Group 1 as most of the European countries, United States, Canada, Saudi Arabia, Oman, Japan and Australia.⁴ These countries show the highest wages and associated standard of

⁴ Noticeable exceptions are Portugal, Poland, Ukraine, and Greece.

living, which can be sustained through businesses that compete with new and unique products and companies that compete through innovation and the production of new and different goods using the most sophisticated production processes.

The interaction between the medium levels and trends in self-employment productivity and high rates of unemployment plays a significant role in forming Group 2. In this group, we find the remaining European countries and some developing and emerging economies, such as southern Africa, Mexico, Brazil, Argentina, Chile, Uruguay, Algeria, Egypt, Turkey, Iran, Kazakhstan, South Korea and Malaysia. These countries produce standard products and services and are susceptible to external, sector-specific demand shocks.

Finally, the countries with the lowest entrepreneurship productivity and the flat-test productivity trend appear in Group 3. In this group, we find Central and Middle African countries, such as Angola, Cameroon, the Central African Republic, Chad, and Nigeria; some Asian countries, such as India, Pakistan, Indonesia and Mongolia; and some South American countries, such as Paraguay, Bolivia, Peru, Ecuador, and Colombia. These are less developed economies, showing limitations in the accessibility of digital technologies, in the level of wages and in competitive advantages accompanied by a heavy reliance on unskilled labor and natural resources.⁵

3.2 Connection with the literature

Our empirical findings call for the leveraging of the most closely related scholarly knowledge on international self-employment development. First, our model-based clustering procedure split the dataset of countries into three distinct groups according to their levels and trends in self-employment productivity. According to the classification of competitiveness across stages of economic development as advocated by Porter et al. (2002), the high-productivity group aligns with economies in the innovation-driven stage, the medium-productivity group relates to efficiency-driven economies, and the low-productivity group contains factor-driven economies.

Second, our results point to the fact that digitalization seems to be a key factor favoring the transition from factor- and efficiency-driven economies to innovation-driven economies. This is likely because digitalization is a key competitive factor, both for a managed economy in which competitiveness is based on efficiency and for capturing the best profit opportunities that favor technological and economic leadership. This agrees with recent findings examining the intersection of digital technologies and entrepreneurship and its impact on the pursuit of sustainable development. In this context, Nambisan (2017) and Jafari-Sadeghi et al. (2021) are two significant examples.

Third, our results suggest that there is a relationship between labor market dynamics (the reduction of unemployment) and the likelihood of a country moving from a medium to a high-productivity self-employment group. At this point, we could argue that a well-functioning labor market generates sufficient wage

⁵ One significant exception is China.

employment opportunities to substantially reduce the relative weight of "necessity entrepreneurs," usually marginally attached to self-employment, with respect to "opportunity entrepreneurs." This results in an increase in the quality of entrepreneurship, becomes a key element in capturing more and better profit opportunities and transforms an economy into an entrepreneurial-driven economy.

Thus, it follows that labor market reforms that promote employability and the provisioning of employment opportunities in the context of a labor market with adequate dynamism are elements that favor the transition to the innovation-driven group. These arguments are in line with the contributions of Acs (2006), Baptista and Thurik (2007), Baumol and Strom (2007), Acs et al. (2008) and Van der Zwan et al. (2016).

Fourth, our statistical evidence does not seem to support the idea that the industrial sector plays a significant role in the probability of a country belonging to the high-productivity self-employment group. In contrast to this result and in line with Lucas (1988), industrialization processes should be accompanied by an increase in self-employment productivity as low-skilled self-employment moves to paid employment as attracted by larger wages in routine industrial job opportunities.

However, our result agrees with the literature on employer-size wage differentials. Davis and Haltiwanger (1996) pointed out that larger employers do not necessarily pay substantially higher wages because the dispersion of wages exhibits a pronounced relationship to employer size. In this context, Poschke (2018) found that it is not only the average size of firms but also their dispersion that is significantly higher in developed countries, and Shi et al. (2020) recently suggested the wage-boosting effect of innovation in shaping firm wages, which does not necessarily depend on firm size. In addition, Acs and Naudé (2013) recognized the complexity of the role of entrepreneurs in industrialization, as this role can be inhibited by, for example, market failures. For this reason, these authors do not see industrial policies as merely functional policies without consideration for firm or entrepreneurial specifics.

Finally, our results do not support that strict employment protection legislation promotes self-employment productivity because it does not influence a transition toward the group of countries with the highest self-employment productivity. This is closely related to the work of Robson (2003), who found very limited evidence for a positive relationship between self-employment and the strictness of employment protection legislation, as this largely depends on the introduction of suitable control variables. Torrini (2005) also failed to find any robust relationship between the self-employment rate and employment protection legislation in a multivariate context.

3.3 Policy recommendations

This study provides important guidelines that policymakers are invited to use when drawing up effective national strategies and policy aspects for self-employment productivity and combatting traditional stereotypes that appear to be less effective for transitioning into the group of highly productive countries.

Our analysis identifies three clusters of countries with some degree of similarity regarding their level and trend in self-employment productivity. This classification can

help national policymakers to verify which group their own country belongs to and determine whether their country has performed on par with other countries in similar economic circumstances.

Unfortunately, our results provide evidence that the catch-up effect, which predicts that all economies will eventually converge in terms of self-employment productivity, does not apply. In this context, we consider the inactivity of policy makers to not be justified and recognize that there is pressure on governments to provide resources to assist in promoting changes toward more productive clusters.

According to our results, the first challenge of policy interventions implies improving incentive structures for entrepreneurs associated with digitalization and promoting the introduction of a new culture of digital entrepreneurship. This implies supporting the development of digital and entrepreneurship skills by addressing some key barriers with a range of policy actions. Examples include embedding digital entrepreneurship modules in entrepreneurship education, offering tailored digital entrepreneurship training programs and improving access to finance for digital entrepreneurship for under-represented and disadvantaged groups.

The second challenge of policy interventions that are aimed at encouraging self-employment productivity to facilitate transitioning into the innovation-driven group requires decisive measures aimed at reducing national unemployment. To name a few, we suggest increasing the attractiveness to private capital, removing the obstacles to labor mobility, improving the quality of formal job allocation mechanisms and reducing rigidities in the housing market. The effectiveness of national policies must also be enhanced by ensuring that funds devoted to reducing unemployment are well managed and that the monitoring and evaluation procedures are improved to guarantee a consistent long-term strategy for human capital.

Likewise, our results point to avoiding policy measures that are likely to be inefficient in promoting self-employment productivity. Our results do not support implementing industrial policies as merely functional policies without consideration of firm or entrepreneurial specifics. In fact, the patterns we find suggest that more numerous industry does not necessarily appear to be better for self-employment productivity.

In addition, our results suggest that changes in the rigidity of labor legislation per se do not serve to stimulate self-employment productivity. Despite the explicit set of rules that govern national employment protection legislation, different degrees of regulatory compliance and the possibility of evasion opportunities could explain this finding. In any case, we consider that the design of incentive schemes should not result in distortions to the allocation of talent between salaried employment and self-employment.

4 Conclusions

This paper reexamines the diversity in the level and dynamics of entrepreneurship across countries in terms of self-employment productivity. To this end, we applied a Bayesian finite mixture model for clustering time series to a large dataset consisting of internationally comparable indicators covering a large set of 121 countries over the last three decades.

Our empirical findings point to the existence of three homogeneous groups stratified by the following levels of entrepreneurship productivity: high-, medium-, and low-productivity countries. These clusters are roughly aligned with the three major groups of countries usually considered in the entrepreneurship literature, namely, factor-, efficiency-, and innovation-knowledge- driven countries, and with the literature on managed vs. entrepreneurial societies. In addition, these clusters parallel the three different stages of economic development, namely, developing, transitioning and developed countries.

In contrast to simpler clustering methods, our clustering approach allows us to examine the key structural variables that allocate each country to a particular cluster and regulate the transition to higher productive groups. In other words, our results not only provide homogeneous country groups regarding the quality of entrepreneurship but also point to the institutions or elements in the national entrepreneurial ecosystem that enable their high self-employment productivity and determine their transition to higher productivity groups. The identification of these factors might be particularly useful for policymakers interested in promoting self-employment productivity and for academics who strive to test previous theories and hypotheses.

Among the factors guiding the transition between groups, we consider some structural variables that strengthen or weaken the probability of a nation becoming an entrepreneurial economy. These variables are the unemployment rate, which refers to the labor market situation; the average of industry added value as a percentage of GDP, which measures the level of industrialization; the Digital Adoption Index, which measures the diffusion and adoption of digital technologies; and the Labor Market Rigidity Index, which is a measure of labor market rigidities. These variables are usually viewed as common drivers in entrepreneurship (e.g., Audretsch and Thurik 2004) and are included in the literature on regional innovation systems and entrepreneurial ecosystems (see Cao and Shi 2021 and Qian and Acs 2023 for recent surveys).

Our results suggest that policy measures oriented toward (i) creating and enabling environments that foster the accessibility of digital technologies and (ii) promoting initiatives for reducing unemployment are key elements for those countries that are generally moving toward becoming highly productive economies. However, the results fail to find the share of industrial added value as a determinant of such transitions. In addition, we find that deregulation policies meant to reduce rigidities in the labor market are also not important keys for transitioning between clusters.

According to our results, the proposed framework is a very promising tool for analyzing the determinants of self-employment productivity. In fact, we look forward to future work addressing the following issues. First, although we focused on aggregate self-employment productivity, we see a natural extension to be the exploration of disaggregated measures, mixed incomes, and nonagricultural self-employment. Second, we could extend the number of additional factors driving the transition between groups. Third, the method is suitable for exploring the determinants of self-employment productivity at the regional level. These extensions were not pursued in this paper due to the cost of reducing the number of observational units. For this reason, these extensions have been explicitly left for further research.

Appendix

Table 4 List of countries and variables values

Country	Code	GDP/SE	U	Industry	LAMRIG	DAI
Angola	AGO	24.497	5.25	55.67	2.42	0.33
Albania	ALB	36.858	16.58	21.99	1.42	0.58
Argentina	ARG	198.993	10.96	26.50	1.55	0.66
Armenia	ARM	45.961	12.36	25.83	1.77	0.61
Australia	AUS	459.129	6.57	25.25	0.98	0.70
Austria	AUT	752.781	4.86	27.10	1.52	0.84
Azerbaijan	AZE	27.612	6.60	47.63	1.72	0.57
Burundi	BDI	2.366	1.59	15.32	1.81	0.25
Belgium	BEL	720.375	7.78	22.17	1.57	0.77
Benin	BEN	7.359	1.33	16.82	1.81	0.22
Burkina Faso	BFA	4.517	3.35	22.44	1.60	0.23
Bangladesh	BGD	11.061	3.65	24.29	1.90	0.34
Bulgaria	BGR	283.491	11.31	24.97	1.73	0.60
Bahrain	BHR	2578.197	1.12	43.11	1.20	0.77
Belarus	BLR	622.949	9.00	35.22	2.47	0.56
Bolivia	BOL	21.674	2.72	27.48	1.86	0.46
Brazil	BRA	89.955	8.57	23.63	2.37	0.67
Botswana	BWA	148.437	19.47	40.53	1.16	0.47
Central African Republic	CAF	2.930	5.65	22.95	1.76	0.14
Canada	CAN	551.234	7.79	27.08	0.79	0.68
Switzerland	CHE	701.452	3.90	26.33	1.26	0.81
Chile	CHL	163.010	7.79	33.61	1.48	0.74
China	CHN	22.342	3.94	44.62	1.65	0.54
Cameroon	CMR	9.267	5.46	27.04	1.89	0.28
Republic of the Congo	COG	19.227	20.11	54.73	1.70	0.30
Colombia	COL	51.576	11.28	29.39	1.89	0.62
Comoros	COM	18.962	5.97	11.61	2.30	0.24
Costa Rica	CRI	137.830	6.79	23.80	2.16	0.63
Czech Republic	CZE	414.745	5.56	33.71	1.52	0.71
Germany	DEU	886.821	7.27	27.84	2.15	0.82
Denmark	DNK	1084.606	6.15	21.61	1.72	0.78
Dominican Republic	DOM	67.367	6.32	30.55	1.51	0.48
Algeria	DZA	121.042	18.29	37.97	0.75	0.40
Ecuador	ECU	49.977	4.20	31.70	1.91	0.55
Egypt	EGY	84.688	10.23	33.60	1.68	0.52
Spain	ESP	471.988	17.14	24.67	2.46	0.75
Ethiopia	ETH	2.864	2.77	12.67	1.39	0.25
Finland	FIN	667.819	10.17	27.35	2.21	0.80
France	FRA	832.679	9.78	19.87	2.23	0.75
Gabon	GAB	171.838	18.70	51.55	1.41	0.36

Table 4 (continued)

Country	Code	GDP/SE	U	Industry	LAMRIG	DAI
United Kingdom	GBR	615.601	6.46	21.16	0.86	0.75
Georgia	GEO	29.714	13.32	21.27	1.77	0.58
Ghana	GHA	11.240	6.26	24.98	1.13	0.42
Guinea	GIN	6.051	4.68	30.34	1.55	0.21
Gambia	GMB	10.023	9.23	14.51	1.29	0.34
Greece	GRC	207.156	13.64	16.87	1.56	0.59
Guatemala	GTM	41.541	2.79	23.72	1.41	0.48
Guyana	GUY	73.140	12.10	22.78	1.27	0.34
Honduras	HND	23.981	4.45	27.69	1.01	0.42
Haiti	HTI	11.292	11.76	26.13	1.17	0.25
Hungary	HUN	433.428	7.94	26.15	1.66	0.66
Indonesia	IDN	28.573	5.25	43.24	1.61	0.42
India	IND	12.873	5.55	28.10	1.38	0.48
Ireland	IRL	675.972	9.29	30.39	0.95	0.65
Islamic Republic of Iran	IRN	95.348	11.25	40.86	1.94	0.46
Italy	ITA	417.851	9.97	23.55	1.96	0.75
Jamaica	JAM	61.279	13.38	21.48	1.16	0.47
Jordan	JOR	318.841	14.79	24.62	1.56	0.54
Japan	JPN	531.854	3.83	30.06	0.50	0.83
Kazakhstan	KAZ	108.193	7.29	34.80	2.08	0.65
Kenya	KEN	13.590	3.09	16.99	1.25	0.43
Kyrgyzstan	KGZ	22.334	7.34	24.19	1.92	0.46
Korea, Republic of	KOR	188.156	3.36	34.45	1.38	0.85
Lebanon	LBN	147.666	8.53	18.03	1.20	0.55
Sri Lanka	LKA	49.667	7.69	27.79	1.61	0.45
Lesotho	LSO	10.967	30.80	32.84	1.29	0.28
Luxembourg	LUX	2649.664	4.07	14.01	2.00	0.85
Morocco	MAR	33.328	11.08	25.97	1.31	0.54
Madagascar	MDG	3.871	3.73	19.30	1.99	0.24
Mexico	MEX	126.083	4.02	32.14	2.01	0.57
North Macedonia	MKD	157.682	30.69	23.47	1.65	0.54
Mali	MLI	6.480	6.64	20.03	1.70	0.30
Mongolia	MNG	33.003	6.01	30.55	1.39	0.53
Mozambique	MOZ	2.208	3.04	18.01	2.21	0.27
Mauritania	MRT	32.243	9.94	30.22	1.82	0.32
Mauritius	MUS	174.592	8.32	23.78	1.21	0.58
Malawi	MWI	5.032	5.80	17.26	1.75	0.25
Malaysia	MYS	171.331	3.32	42.48	0.85	0.67
Namibia	NAM	85.788	21.00	26.15	1.06	0.38
Niger	NER	2.931	1.42	20.87	1.66	0.16
Nigeria	NGA	15.819	4.53	28.11	1.27	0.39
Nicaragua	NIC	25.185	6.26	22.12	0.82	0.42

Table 4 (continued)

Country	Code	GDP/SE	U	Industry	LAMRIG	DAI
Netherlands	NLD	718.543	5.08	21.10	2.16	0.83
Norway	NOR	1484.825	4.09	32.79	2.07	0.79
Nepal	NPL	5.829	2.02	16.65	1.90	0.33
New Zealand	NZL	392.084	6.07	23.09	0.49	0.69
Oman	OMN	1454.481	3.79	55.44	1.30	0.65
Pakistan	PAK	19.697	1.30	20.14	1.15	0.38
Panama	PAN	134.825	3.58	21.73	2.42	0.56
Peru	PER	31.115	4.17	31.43	1.65	0.54
Philippines	PHL	32.266	3.48	33.38	1.64	0.47
Papua New Guinea	PNG	11.934	2.55	34.23	1.01	0.32
Poland	POL	202.436	11.70	29.27	2.25	0.67
Portugal	PRT	277.398	7.92	21.57	2.45	0.76
Paraguay	PRY	46.487	5.43	35.19	1.71	0.50
Russian Federation	RUS	1023.615	7.43	32.82	2.25	0.72
Rwanda	RWA	3.422	0.87	17.32	1.77	0.42
Saudi Arabia	SAU	2304.870	5.68	53.52	1.02	0.67
Sudan	SDN	34.608	15.86	18.33	1.67	0.29
Senegal	SEN	14.330	6.84	23.52	1.66	0.34
Singapore	SGP	847.965	3.99	28.64	0.86	0.87
Sierra Leone	SLE	4.444	3.89	17.04	1.81	0.25
El Salvador	SLV	45.943	6.12	25.64	1.77	0.49
Suriname	SUR	258.692	10.10	29.41	0.85	0.46
Sweden	SWE	844.297	7.26	24.34	2.33	0.82
Chad	TCD	4.402	0.89	14.21	1.83	0.20
Togo	TGO	6.282	3.44	18.27	1.73	0.23
Thailand	THA	39.700	1.32	37.20	1.77	0.59
Trinidad and Tobago	TTO	227.752	9.18	47.84	0.58	0.55
Tunisia	TUN	107.383	14.97	27.23	1.72	0.54
Turkey	TUR	143.462	9.30	27.84	1.72	0.61
Tanzania	TZA	4.288	2.97	21.21	1.63	0.32
Uganda	UGA	5.429	2.39	21.15	1.71	0.31
Ukraine	UKR	177.219	7.73	30.41	2.23	0.49
Uruguay	URY	138.175	9.79	24.10	1.27	0.74
United States	USA	1496.202	5.85	20.31	0.65	0.73
Uzbekistan	UZB	22.841	7.07	24.81	1.55	0.36
Vietnam	VNM	11.576	1.90	33.85	1.75	0.49
South Africa	ZAF	268.835	28.02	27.15	1.16	0.61
Zambia	ZMB	9.058	13.18	31.89	1.23	0.32
Zimbabwe	ZWE	12.636	5.17	25.76	0.86	0.32

GDP/SE is the average GDP in thousands (constant international 2017 dollars, PPP) per self-employed person; *U* unemployment rate; Industry is value added by industry as a percentage of GDP; *LAMRIG* Labor Market Rigidity Index, *DAI* Digital Adoption Index

Acknowledgements The authors would like to thank Jesús Crespo-Cuaresma, Begoña Cueto, Frank Fossen, Javier J. Pérez, Concepción Román, and Andre Van Stel for their insightful comments that contributed substantially to the development of this paper. All remaining errors are our responsibility.

Authors' contributions All authors contributed in the same way to the writing and preparation of this research article.

Funding Funding for open access publishing: Universidad de Huelva/CBUA. The authors acknowledge funding from the Spanish Ministry of Science, Innovation and Universities through project PID2020-115183RB-C22; from *Junta de Andalucía* through grant P20-00733 and Research Group SEJ-487 (Spanish Entrepreneurship Research Group – SERG); and from Research and Transfer Policy Strategy of the University of Huelva 2021. M. Camacho is grateful for the support of grant PID2022-136547NB-I00 funded by MICIU/AEI/<https://doi.org/10.13039/501100011033> and by FEDER, UE.

Data availability The data and MATLAB code that support the findings of this study are available from the corresponding author upon request.

Declarations

Competing interests The authors have no competing interests to declare.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Acs ZJ (2006) How is entrepreneurship good for economic growth? *Innovations: Techno Govern Glob* 1(1):97–107
- Acs ZJ, Naudé W (2013) Entrepreneurship, stages of development, and industrialization. In: Szirmai A, Naudé W, Alcorta L (eds) *Pathways to industrialization in the twenty-first century: New challenges and emerging paradigms*. Oxford Academic, Oxford
- Acs ZJ, Desai S, Hessels J (2008) Entrepreneurship, economic development and institutions. *Small Bus Econ* 31(3):219–234
- Acs ZJ, Audretsch DB, Evans DS (1994) Why does the self-employment rate vary across countries and over time? Discussion Paper No. 871, Center for Economic Policy Research, January 1994. <https://econpapers.repec.org/paper/cprceprdp/871.htm>
- Allub L, Erosa A (2019) Financial frictions, occupational choice and economic inequality. *J Monet Econ* 107:63–76
- Anokhin S, Schulze WS (2009) Entrepreneurship, innovation, and corruption. *J Bus Ventur* 24(5):465–476
- Arin KP, Huang VZ, Minniti M, Nandialath AM, Reich OFM (2015) Revisiting the determinants of entrepreneurship: a Bayesian approach. *J Manag* 41(2):607–631
- Audretsch DB, Thurik AR (2004) A model of the entrepreneurial economy. *Int J Entrep Educ* 2(2):143–166
- Baptista R, Thurik AR (2007) The relationship between entrepreneurship and unemployment: is Portugal an outlier? *Technol Forecast Soc Change* 74(1):75–89

- Baumol WJ, Strom RJ (2007) Entrepreneurship and economic growth. *Strateg Entrep J* 1(3–4):233–237
- Belitski M, Chowdhury F, Desai S (2016) Taxes, corruption, and entry. *Small Bus Econ* 47(1):201–216
- Bjørnskov C, Foss NJ (2016) Institutions, entrepreneurship, and economic growth: what do we know and what do we still need to know? *Acad Manag Perspect* 30(3):292–315
- Blanchflower DG (2000) Self-employment in OECD countries. *Labour Econ* 7(5):471–505
- Blanchflower DG (2004) Self-employment: More may not be better. *Swed Econ Policy Rev* 11(2):15–73
- Blanchflower DG, Shadforth C (2007) Entrepreneurship in the UK. *Found Trends Entrep* 3(4):257–364
- Campos NF, Nugent JB (2018) The dynamics of the regulation of labour in developing and developed countries since 1960. In: Campos NF, Grauwe P, Ji Y (eds) *The political economy of structural reforms in Europe*. Oxford University Press, pp 75–88
- Cao Z, Shi X (2021) A systematic literature review of entrepreneurial ecosystems in advanced and emerging economies. *Small Bus Econ* 57(1):75–110
- Centeno M (2000) Is self-employment a response to labour market rigidity? *Econ Bull* 37–44
- Congregado E, Golpe AA, Carmona M (2010) Is it a good policy to promote self-employment for job creation? Evidence from Spain. *J Policy Model* 32(6):828–842
- Cowling ML, Wooden M (2021) Does solo self-employment serve as a ‘stepping stone’ to employership? *Labour Econ* 68:101942
- Davis SJ, Haltiwanger J (1996) Employer size and the wage structure in US manufacturing. *Ann Econ Stat* 41–42:323–367
- Djankov S, Ganser T, McLiesh C, Ramalho R, Shleifer A (2010) The effect of corporate taxes on investment and entrepreneurship. *Am Econ J: Macroeconomics* 2(3):31–64
- Dutta N, Sobel R (2016) Does corruption ever help entrepreneurship? *Small Bus Econ* 47:179–199
- Estrin S, Korosteleva J, Mickiewicz T (2012) Which institutions encourage entrepreneurial growth aspirations? *J Bus Ventur* 28(4):564–580
- Fairlie RW, Fossen FM (2020) Defining opportunity versus necessity entrepreneurship: two components of business creation. In: Polachek SW, Tatsiramos K (eds) *Change at home, in the labor market, and on the job*, vol 48. Emerald Publishing Limited, pp 253–289. <https://doi.org/10.1108/S0147-9122020000048008>
- Fölster S (2002) Do lower taxes stimulate self-employment? *Small Bus Econ* 19:135–145
- Frühwirth-Schnatter S (2006) *Finite mixture and Markov switching models*. Springer, New York
- Frühwirth-Schnatter S, Kaufmann S (2008) Model-based clustering of multiple time series. *J Bus Econ Stat* 26(1):78–89
- Hamilton JD, Owyang MT (2012) The propagation of regional recessions. *Rev Econ Stat* 94(4):935–947
- Jafari-Sadeghi V, Garcia-Perez A, Candelo E, Couturier J (2021) Exploring the impact of digital transformation on technology entrepreneurship and technological market expansion: the role of technology readiness, exploration and exploitation. *J Bus Res* 124:100–111
- Kannianen V, Vesala T (2005) Entrepreneurship and labor market institutions. *Econ Modell* 22(5):828–847
- Kaufmann S (2010) Dating and forecasting turning points by Bayesian clustering with dynamic structure: a suggestion with an application to Austrian data. *J Appl Economet* 25(2):309–344
- Lucas RE Jr (1988) On the mechanics of economic development. *J Monet Econ* 22(1):3–42
- Maloney WF (2004) Informality revisited. *World Dev* 32(7):1159–1178
- Nambisan S (2017) Digital entrepreneurship: Toward a digital technology perspective of entrepreneurship. *Entrepre Theory Pract* 41(6):1029–1055
- Pietrobelli C, Rabellotti R, Aquilina M (2004) An empirical study of the determinants of self-employment in developing countries. *J Int Dev* 16(6):803–820
- Porter M, Sachs J, McArthur J (2002) Executive summary: Competitiveness and stages of economic development. In: Porter M, Sachs J, Cornelius PK, McArthur J, Schwab K (eds) *The global competitiveness report*. Oxford University Press, New York, pp 2001–2002
- Poschke M (2013) Who becomes an entrepreneur? Labor market prospects and occupational choice. *J Econ Dyn Control* 37(3):693–710
- Poschke M (2018) The firm size distribution across countries and skill-biased change in entrepreneurial technology. *Am Econ J: Macroeconomics* 10(3):1–41
- Poschke M (2019) Wage employment, unemployment and self-employment across countries. IZA Discussion Paper No. 16271. <https://doi.org/10.2139/ssrn.3401135>
- Qian H, Acs ZJ (2023) Entrepreneurial ecosystems and economic development policy. *Econ Dev Q* 37(1):96–102

- Robson MT (2003) Does stricter employment protection legislation promote self-employment? *Small Bus Econ* 21(3):309–319
- Rodriguez-Santiago A (2022) Re-evaluating the relationship between economic development and self-employment, at the macro-level: a Bayesian model averaging approach. *Int J Interact Multimedia Artif Intell* 7(3):20–25
- Scott SL (2011) Data augmentation, frequentist estimation, and the Bayesian analysis of multinomial logit models. *Stat Pap* 52(1):87–109
- Shane S (2009) Why encouraging more people to become entrepreneurs is bad public policy. *Small Bus Econ* 33:141–149
- Shi L, Li S, Fu X (2020) The fourth industrial revolution, technological innovation and firm wages: Firm-level evidence from OECD economies. *Revue d'Économie Industrielle* 1(169):89–125
- Sobel RS (2008) Testing Baumol: Institutional quality and the productivity of entrepreneurship. *J Bus Ventur* 23(6):641–655
- Stam E (2015) Entrepreneurial ecosystems and regional policy: A sympathetic critique. *Eur Plan Stud* 23(9):1759–1769
- Stam E, Van de Ven A (2021) Entrepreneurial ecosystem elements. *Small Bus Econ* 56(2):809–832
- Thurik AR, Carree MA, Van Stel A, Audretsch DB (2008) Does self-employment reduce unemployment? *J Bus Ventur* 23(6):673–686
- Torrini R (2005) Cross-country differences in self-employment rates: The role of institutions. *Labour Econ* 12(5):661–683
- Urbano D, Audretsch D, Aparicio S, Noguera M (2020) Does entrepreneurial activity matter for economic growth in developing countries? The role of the institutional environment. *Int Entrepre Manag J* 16(3):1065–1099
- Van der Zwan P, Thurik AR, Verheul I, Hessels J (2016) Factors influencing the entrepreneurial engagement of opportunity and necessity entrepreneurs. *Eurasian Bus Rev* 6(3):273–295
- Wennekers S, Van Stel A, Carree M, Thurik AR (2010) The relationship between entrepreneurship and economic development: Is it U-shaped? *Found Trends Entrepre* 6(3):167–237

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.