

Re-Evaluating the Relationship Between Economic Development and Self-Employment, at the Macro-Level: A Bayesian Model Averaging Approach

Ana Rodriguez-Santiago*

Vienna University of Economics and Business (Austria), University of Huelva (Spain) & CCTH - Centro Científico Tecnológico Huelva (Spain)

Received 21 June 2021 | Accepted 24 January 2022 | Published 26 January 2022



ABSTRACT

We re-evaluate the relationship between stages of economic development and entrepreneurship, at the macro level. We first conduct a literature review of previous empirical research on cross-country determinants of entrepreneurship in order to put our contribution in perspective. To circumvent problems related to model uncertainty we use Bayesian Model Averaging (BMA) to evaluate the robustness of determinants of economic growth in a new dataset of 117 countries in the 2005-2019 period, allowing fixed effects and investigating the existence of heterogeneity allowing interactions of our focus variable with other regressors. Our empirical analysis then shows that the variation of self-employment rates across countries are mainly determined by variations in the unemployment, the stage of economic development and the variations in labor market frictions. When interactions are taken into account, results confirm that there is a differential effect of labor market frictions in countries with different levels of income. Frictions in labor market may encourage becoming self-employed in richer countries.

KEYWORDS

Bayesian Model Averaging, Entrepreneurship, Macro-level Determinants, Model Uncertainty, Panel, Self-employment.

DOI: 10.9781/ijimai.2022.01.007

I. INTRODUCTION

THE empirical literature on the macro-level determinants of entrepreneurship/self-employment¹ has analyzed a wider set of predictors as potential entrepreneurship drivers. These potential determinants relate to human capital², the level of development³ and institutions⁴. There is a great number of studies in which a large set of regressors are included in so-called ‘ad-hoc’ regressions, based on previous hypotheses and theoretical propositions.⁵

Whatever the type of specification is -structural or not⁶ - and independently of the inclusion of a focus variable, we have a set of theories and propositions not mutually exclusive and, as in other fields of economics research, most of the empirical results in previous literature on the determinants of entrepreneurship at the macro-level have potential problems of model uncertainty, that is, regarding the choice of predictors.

To the best of our knowledge, we only can find two previous attempts to circumvent these potential problems. On the one hand, [16] adopted an algorithmic approach based on resampling and bootstrap techniques in a cross section of 69 countries for the year 2014, using data drawn from the Global Entrepreneurship Monitoring Database (henceforth, GEM). In short, the method is a step-by-step approach for finding the subset of explanatory variables leading the best possible prediction accuracy. With this strategy they select the more relevant regressors for explaining the national total entrepreneurship activity (TEA). The strength of Innovation and research and the level of entrepreneurial education are the best predictors in their analysis. [17] adopted an alternative solution. They applied a Bayesian model averaging (henceforth, BMA) to address the issue of model uncertainty in the framework of the literature on the determinants of self-employment, following the seminal contribution of [19], who combined the Bayesian Information criteria model weights and

¹ Throughout this paper, we use the terms entrepreneurship and self-employment synonymously and interchangeably. This operationalization of entrepreneurship as self-employment is dictated by data availability considerations.

² Educational attainment and sociodemographic characteristics.

³ Economic development, macroeconomic stability -unemployment, inflation, government size-, financial development and access to finance and technological progress.

⁴ Labor market institutions, Globalization, Administrative complexity and the rule of law, Taxes and Government.

⁵ These works may be classified into two groups: with or without focus variable. For example, among the former are the works of [1]-[14] and among the latter the works of [15]-[17].

⁶ The adjective structural describes how the specification is derived from a theoretical model. As [18] states, this approach allows to understand how the model is identified. The works of [3], [4], [13] and [14] are examples of this approach in the empirical literature on the determinants of entrepreneurship.

* Corresponding author.

E-mail address: ana.rodriguez@dege.uhu.es

maximum likelihood estimates for model selection, later revisited in the works of [20] and [21]. By using 32 predictors, aggregated into three groups – human capital, level of development and Institutions–, they use the BMA approach for correcting model uncertainty. With a short panel of 80 observations drawn from the GEM, the gross domestic product per capita, the unemployment and tax rates and the volatility of inflation are identified as the best predictors of the entrepreneurship rate, when model uncertainty is corrected for.

Despite the advantages of this last approach, the poor quality of the database and short period of observation and the non-consideration of interactions awake serious concerns about the robustness of the last two previous contributions. The problem may be particularly worrying if the relationship between self-employment and the potential regressors was dependent on the state of economic development, as suggested several previous contributions [22]-[25].

The present study aims to re-evaluate the robustness of the statistical significance of 21 macrolevel variables as predictors of the cross-country differences in the level of self-employment taking into account the potential parameter heterogeneity according to country development level. To this end, we use an extension of the BMA, suggested by [26], to re-evaluate the robustness of 21 determinants of self-employment in a new larger dataset of 117 countries during the period from 2005 to 2019, and investigate the existence of parameter heterogeneity allowing interactions between potential regressors and the stage of economic development based in panel data with fixed effects.

This article contributes to the previous empirical literature on self-employment determinants on the following grounds.

First, we provide new (and updated) empirical evidence on the drivers of self-employment in a much larger dataset than in the available empirical literature, including both developed and developing countries. As usual in prior related literature joint to our focus variable –the economic development proxied by GDP per capita–, a set of control variables are also included –e.g., proxies of different type of institutions, human capital, openness and technological progress, among others–.

Second, although previous empirical literature devoted to the identification the drivers of entrepreneurship across countries is considerable, there is a lack of consensus. Empirical evidence has not provided unambiguous results and as a result some controversies, about what are the drivers (and barriers) of entrepreneurship, have emerged, with deep policy implications. These inconsistencies may be due to the poor quality of data, to problems related with measurement issues of some variables and to the discretionary choice of predictors, the so-called model uncertainty [27].⁷ To circumvent this problem of specification we use an extended version of the BMA for panel data allowing interactions and parameter heterogeneity [28] and [26] in which inference is based on a weighted average of all possible model specifications, not in a particular one. To the best of our knowledge this contribution is the first attempt to use the BMA approach with interactions in the context of the empirical literature on the determinants of entrepreneurship/self-employment.

Third, the data collected in the new database have been drawn from different sources –International Labor Organization Statistics, OECD Statistics, Penn World Tables (10.0), World Bank and World Intellectual Property Organization–. For measuring some explanatory variables, alternative indicators were taken into consideration to enlarge the sample.

Empirical support is found for the view that national self-employment rate is affected by unemployment, labor market frictions and the level of economic development –a nonlinear relationship

consistent with the observed U-shape relationship between GDP and self-employment–. When interactions are considered, the key finding is that labor market frictions for the most advanced countries economic are found to be associated with higher self-employment rates.

The paper proceeds as follows. In Section II we conduct a brief description of the methodology that we employ and data. Section III describes the empirical results and, finally, Section IV concludes.

II. METHODS AND DATA

A. Data

Our sample consists of a balanced panel data set formed by 117 economies over the period 2005–2019. Entrepreneurship is operationalized in terms of self-employment, reflecting data availability at the time-series level⁸. Entrepreneurship is defined as the self-employment rate, which is the number of business owners –employers and solo self-employed workers– divided by the total labor force.⁹ The self-employment rate is drawn from the International Labor Organization Statistics (ILO-Statistics).

To explain the cross-national variations on self-employment rate, we include the 21 following variables (see Table A.II in the appendix for sources and descriptive statistics):

GDP per capita on purchasing power parity (PPP): gross domestic product converted to international dollars using purchasing power parity rates. Data are in constant 2017 international dollars.

Agriculture, Services and Industry correspond to the ISIC divisions 1-5, 50-99 and 10-45, respectively.

Exports/Imports of goods and services represent the value of all goods and other market services provided/received to/from the rest of the world.

Patent applications are worldwide patent applications filed through the Patent Cooperation Treaty procedure or with a national patent office for exclusive rights for an invention.

Internet users. This indicator captures the proportion of individuals using the Internet based on results from national household surveys.

Human capital index. Index provided by the Penn World Tables based on the average years of schooling and an assumed rate of return to education, based on Mincer equation estimates around the world.

Female Labor force participation rate. Proportion of females aged 15 and older who are economically active.

Frictions in Labor Markets. Following [13] we use the unemployment-wage employment ratio as an indicator of labor market frictions. He argues that labor market frictions make it more difficult to find a job and cause high levels of unemployment relative to wage employment, reducing the opportunity cost of self-employment.

Unemployment (Youth unemployment). Share the labor force that is without work but available for and seeking employment (in the age interval 15-24, for the younger age group).

Rural population. It refers to people living in rural areas as defined by national statistical offices. It is calculated as the difference between total population and urban population.

Total population. It is “de facto” definition of population, which counts all residents regardless of legal status or citizenship.

⁸ Table A.I and Fig. A.1 in the appendix show a list of the countries included on the sample and a map with the average self-employment rate over the sample period, respectively.

⁹ This is a common practice, for convenience although it is aware that entrepreneurship is a multifaceted concept and look for better indicators is a major challenge for empirical research.

⁷ The potential bias of ignoring this uncertainty is discussed in the works of [22], [23], [26] and [27]. See [29] for a detailed and recent survey.

Inflation. Proxied by the annual growth rate of the GDP implicit deflator.

Taxes. The total tax and contribution rate measures the amount of taxes and mandatory contributions borne by the business in the second year of operation, expressed as a share of commercial profit. The labor tax and contributions measures all government mandated labor contributions that are borne by the business in the second year of operation, expressed as a share of commercial profit.

Doing Business. The score for starting a business is the simple average of the scores for each of the component indicators: the procedures, time and cost for an entrepreneur to start and formally operate a business, as well as the paid-in minimum capital requirement.

Control of Corruption. This index captures perceptions of the extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as “capture” of the state by elites and private interests.

Government Effectiveness. It captures perceptions of the quality of public services, the quality of the civil service and the degree of its independence from political pressures, the quality of policy formulation and implementation, and the credibility of the government’s commitment to such policies.

B. Methodology

As we mentioned, our objective is to select the appropriate specification or statistical model for the determinants of self-employment avoiding the personal discretion of the researcher. Consider the general model,

$$y = \alpha + X_k \beta_k + \varepsilon, \quad \varepsilon \sim N(0, \sigma^2 I) \quad (1)$$

Where y is the self-employment rate and k is the number of regressors included, from all the possible regressors K . We are interested in the effect β of every particular variable and interaction included in X . With 21 possible variables, the cardinality of the model space including interactions would be 241, number of combinations of the 41 variables/interactions in models of size from 1 to 41. It is not possible to estimate around 2.199 billion models. If we could estimate all the models and get the probabilities of every model, the posterior distribution of the parameter β would be a weighting of the estimate of β from every particular model M_i times the probability that this model is true given the data.

$$p(\beta|y) = \sum_{i=1}^{2^K} p(\beta|y, M_i) p(M_i|y) \quad (2)$$

We use a BMA approach, first introduced by [19], to assess the implicit uncertainty across models. With BMA we assign a prior probability to a set of models and update it according to the data. Then, the posterior model probabilities (PMP) of the top models are averaged to calculate the posterior inclusion probabilities (PIP) for the potential determinants.

The PMP of every model is approximated by the marginal likelihood times the prior probability of the model, not conditional on the data.

$$p(M_i|y) \propto p(y|M_i) p(M_i) \quad (3)$$

The researcher is in charge of including the prior beliefs on the model prior. Non-informative prior will assume $p(M_i) = 1/2^K$, assessing the same probability to all the possible models. Under this prior, the posterior model probability will be proportional to the marginal likelihood. It is the likelihood function after integrating away all the parameters of the model (α, β, σ):

$$p(y|M_i) = \iiint p(y|M_i, \alpha, \beta_k, \sigma) p(\alpha, \beta_k, \sigma) d\alpha d\beta_k d\sigma \quad (4)$$

Priors for model-specific parameters. Setting uninformative prior, we let the data speak. We establish non-informative priors on intercept $p(\alpha) \propto 1$ and on the deviation $p(\sigma) \propto 1/\sigma$. But, in order to

find an analytical solution of the marginal likelihood, we need barely informative prior for coefficients β . We assume informative prior on β given σ by the g -prior by [30]

$$p(\beta_k|\sigma) \sim \mathcal{N}(0, \sigma^2(gX'X)^{-1}) \quad (5)$$

This prior requires only elicitation of g . The variance-covariance matrix of β has the same structure of the variance-covariance matrix of OLS estimator, scaled with g , that determines the shrinkage in the regression parameters

$$E(\beta|y, M_i) = \frac{1}{1+g} (X'X)^{-1} X'y = \frac{1}{1+g} \hat{\beta}_{OLS} \quad (6)$$

The marginal likelihood for model M_i is given by

$$p(y|M_i) \propto \left(\frac{g}{1+g}\right)^{\frac{k_i}{2}} \left[\frac{1}{1+g} y'M_x y + \frac{g}{1+g} (y - \bar{y}_n)'(y - \bar{y}_n) \right]^{-\frac{n-1}{2}} \quad (7)$$

with the residual matrix $M_x = (I - X(X'X)^{-1}X')$.

The Bayes factor comparing M_i to the null model is given by

$$BF[M_i : M_0] = \frac{p(y|M_i)}{p(y|M_0)} = \left(1 + \frac{1}{g}\right)^{\frac{n-k_i-1}{2}} \left[1 + \frac{1}{g}(1 - R_i^2)\right]^{-\frac{n-1}{2}} \quad (8)$$

Fixing g , the marginal likelihood depends on how well the model fits the data and the size of the model. The use of the g -prior leads to a marginal likelihood which incorporates Occam’s razor properties: For a given value of k_i , $p(y|M_i)$ and $BF[M_i : M_0]$ increase as goodness of fits increases, and for a given goodness of fit, $p(y|M_i)$ and $BF[M_i : M_0]$ increase as k_i decreases.

Literature has provided different options when choosing g . Unit Information Prior (UIP), proposed by [31], establishes $g = n$, which implies that the Bayes Factor mimics BIC [32]. Risk Inflation Criterion (RIC), proposed by [33], sets $g = K^2$, that minimizes the maximum increase in risk due to selecting rather than knowing the correct predictors. According to [34], we use the Benchmark prior (BRIC), $g = \max(n, K^2)$, that will decide between UIP or BIC depending on the number of potential regressors K and the sample size n .

Priors over the model space. We follow [21] for the specification of the prior model probabilities. We establish a fully random prior for the model and a binomial-beta hyperprior over prior inclusion probability with prior expected model size $\bar{k} = K/2$. This hyper-prior leads to flat prior inclusion probability¹⁰.

Related predictors. In order to know the different determinants of self-employment rate depending on the income level, we include in our model interactions of all the variables with the GDP per capita. Since we want to analyze the determinants of the self-employment comparing different level of development, we need to control by the effect of individual variables to compare the effect of the interactions. Following [26], we include the specification of strong heredity principle based on [35], which is a special case of George’s dilution priors [36]. This way, we define prior probabilities across models where interactions are not present or are present with parent variables, and assign zero prior probability to models with interactions where some parent variable is not present.

The rationale behind this specification is that using a uniform prior over the model space we are interpreting an interaction term as an exclusive effect of that particular product of covariates and ignoring the independent effects of the interacted variables. Since we want to assess the differential effect of the covariates depending on GDP, we need to evaluate the significance of these interactions in a model which contains linear terms in both variables in addition to the interaction variable.

¹⁰ In order to check robustness, we tried an informative specification for expected model size ($\bar{k} = 5$). Results do not show significant change.

Computational Issues. Sampling from the model space. Following [20] we use Markov Chain Monte Carlo Model Composition (MC3) to approximate the posterior model probability. Starting with a random model with a random number of variables, we compute the posterior model probability and then propose a candidate model in the neighborhood of the first model, with one variable more or less, randomly chosen. Then, we can compare the posterior model probability with the previous one and keep the model with a higher value, that will be compared with a new candidate from the neighborhood. This procedure will visit models with higher non-negligible posterior model probability. Convergence of the MC3 sampler can be checked by computing the correlation between analytical and frequency-based posterior model probabilities for a region of the model space. We estimate 5.000.000 draws, discard the first 1.000.000 draws as the burn-in sample, and compute the results based on the top 100 models visited by the Markov chain.

Using the extension of the BMA methodology [20] to a panel data framework [28], we estimate a country fixed effects panel, including interactions terms with GDP per capita under the strong heredity prior over the model space. We present posterior inclusion probabilities (PIP)¹¹, the mean of the posterior distribution for each parameter (and interaction) and the corresponding posterior standard deviation (SD).

III. RESULTS

Table I presents the results of the BMA exercise. We use the benchmark BRIC prior and establishes a binomial-beta prior on a prior expected model size of $K/2 = 20.5$. Using the strong heredity priors, we only evaluate models which contain the parent variables when interaction terms are included.

Fig. 1 shows the variables inclusion of variables with highest PIP on the top visited models and the sign when included. Our analysis, based on 21 covariates and the interaction of GDPpc with all the variables, presents a posterior mean model size of 11 variables but identifies only five variables/terms as significantly determinants of the self-employment.

First, GDP per capita presents a negative and statistically significant relationship with self-employment rate, in line with previous literature [13], [15], [17]. Cross-country analysis show self-employment rates are lower in richer countries [13] while some propositions and theories have attempted to provide a rationale for this negative relationship [37]-[38]. [23] distinguishes three major stages of development in self-employment rates. The first is characterized by high rates of non-agricultural self-employment. The second is characterized for a growing number of transitions to the wage-employment sector. As the economy becomes more developed fewer people become self-employed. In the third, the business sector expanded relative to manufacturing and the improvement in information technologies increase the returns of entrepreneurship. From this perspective, a U-shape relationship between self-employment and economic development emerges. Both arguments suggest a non-linear relationship as the significance of the coefficient associated to the quadratic GDPpc seems to indicate [39].

The next variables appearing as dominant determining the self-employment rates are related to the labor market. Unemployment rate emerges as negative and statistically significant, providing support to the entrepreneurial-pull hypothesis. As [40] states it has been a traditional source of controversy among economists, caused by the two competing hypotheses provided by the theory. The recession-push hypothesis which states that in times of crisis the lack of job opportunities pushes

unemployed into self-employment. By contrast, the prosperity-pull mechanism suggests a positive comovement between self-employment and economic opportunities. If this relationship prevails in times of crisis, entrepreneurs are “pulled” out of self-employment, suggesting the existence of negative relationship between unemployment and self-employment. Previous empirical literature provides a large array of different results. As a result, the exact nature of the relation is still not clear, since we can only aspire to capture the net effect [15], [17]. Our results support the prosperity pull hypothesis.

TABLE I. BMA RESULTS

| Variable | PIP | M | SD |
|--------------------|--------------|------------|-------|
| GDPpc | 1,00 | -40,71*** | 12,23 |
| AGR | 0,05 | 0,03 | 0,17 |
| EXP | 0,48 | -0,01 | 0,01 |
| IMP | 0,53 | -0,01 | 0,02 |
| SER | 0,11 | 0,00 | 0,02 |
| IND | 0,02 | 0,00 | 0,00 |
| PAT | 0,01 | 0,00 | 0,00 |
| INT | 1,00 | -0,03 | 0,03 |
| HUC | 0,08 | -0,40 | 2,15 |
| LFF | 0,56 | 0,04 | 0,05 |
| UWE | 1,00 | -171,20*** | 12,28 |
| UNE | 1,00 | -1,28*** | 0,10 |
| UNY | 0,08 | 0,00 | 0,03 |
| RUR | 0,21 | 0,14 | 0,31 |
| POP | 0,01 | 0,00 | 0,00 |
| INF | 0,02 | 0,00 | 0,00 |
| TTX | 0,05 | 0,00 | 0,01 |
| LTX | 1,00 | 0,03 | 0,17 |
| BUS | 0,70 | -0,01 | 0,01 |
| COR | 0,73 | -5,24 | 4,60 |
| GOV | 0,05 | -0,03 | 0,25 |
| GDPpc ² | 0,95 | 1,70*** | 0,59 |
| GDPpc x AGR | 0,04 | 0,00 | 0,02 |
| GDPpc x EXP | 0,01 | 0,00 | 0,00 |
| GDPpc x IMP | 0,02 | 0,00 | 0,00 |
| GDPpc x SER | 0,00 | 0,00 | 0,00 |
| GDPpc x IND | 0,00 | 0,00 | 0,00 |
| GDPpc x PAT | 0,00 | 0,00 | 0,00 |
| GDPpc x INT | 0,06 | 0,00 | 0,00 |
| GDPpc x HUC | 0,03 | 0,03 | 0,20 |
| GDPpc x LFF | 0,01 | 0,00 | 0,00 |
| GDPpc x UWE | 1,00 | 23,94*** | 1,60 |
| GDPpc x UNE | 0,01 | 0,00 | 0,01 |
| GDPpc x UNY | 0,00 | 0,00 | 0,00 |
| GDPpc x RUR | 0,19 | -0,02 | 0,03 |
| GDPpc x POP | 0,00 | 0,00 | 0,00 |
| GDPpc x INF | 0,00 | 0,00 | 0,00 |
| GDPpc x TTX | 0,01 | 0,00 | 0,00 |
| GDPpc x LTX | 0,17 | 0,01 | 0,02 |
| GDPpc x BUS | 0,01 | 0,00 | 0,00 |
| GDPpc x COR | 0,57 | 0,49 | 0,47 |
| GDPpc x GOV | 0,00 | 0,00 | 0,02 |
| PMS | 11,77 | | |
| Corr. PMP | 0,9998 | | |

PIP, Posterior inclusion probability; M, mean of the posterior distribution parameter; SD, posterior standard deviation of the parameter; PMS, posterior mean model size; PMP, posterior model probability. Statistics based on the 100 most visited models by the Markov chain. Bold entries refer to variables who PIP>0.5. *, p<0.10; **, p<0.05; ***, p<0.01.

Finally, the frictions on the labor market are found to be a determinant of the variation of self-employment across countries. The relationship between the ratio unemployed by wage employees and self-employment is significant and negative. When checking the importance of interaction terms of GDPpc, only the one with the ratio U/WE appears to be significant. It means that economies with more frictions on the labor market tend to present lower self-employment rate, unless they have higher level of development, in which case the relationship between frictions and self-employment turns positive. This outcome is in line with the results provided by [1], [4], [13].

¹¹ PIP is considered robust when higher than the prior inclusion probability (π), which is expected model size by the number of variables. For the flat prior over the model space $\tilde{k} = K/2$, $\pi = \tilde{k}/K = 0.5$.

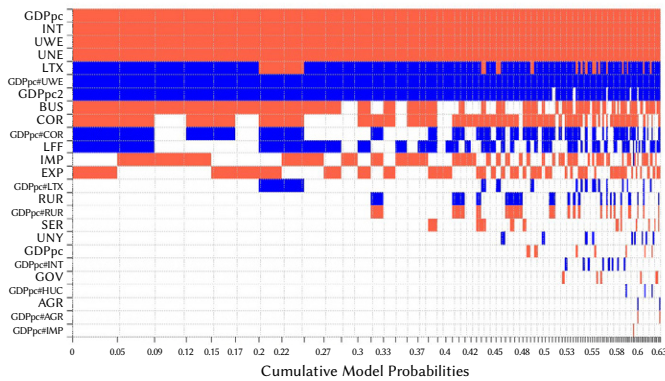


Fig. 1. Selected models probabilities. Inclusion and sign of variables. In blue, positive sign; in red, negative sign; and in white, non-inclusion.

IV. CONCLUSION

The contribution of this paper was to provide empirical evidence on the drivers of self-employment in a new and much larger – and harmonized– dataset than in the available empirical literature including 117 countries observed 20 periods and a set of 21 potential entrepreneurship determinants. As usual in prior related literature, joint to our focus variable –the economic development proxied by GDP per capita– a large battery of control variables is also included –e.g., GDP per capita square, institutions, human capital, openness and technological progress, among others– and data and we include a new proxy for capturing frictions in the labor market suggested by [13]. To circumvent problems associated to model uncertainty we adopted a BMA approach for panel. Our results provide a new explanation of the cross-country differences in the level of self-employment. We show that the unemployment rate, the frictions in the labor market and the stage of economic development are strong determinants of self-employment across the 117 countries included in our sample. Other potential drivers are not significantly correlated with self-employment.

APPENDIX

TABLE A.I. COUNTRIES IN THE SAMPLE

| | | | | |
|---------------|----------------|-------------|--------------|----------------|
| Albania | Cote d'Ivoire | Indonesia | Moldova | Serbia |
| Algeria | Croatia | Iran | Mongolia | Sierra Leone |
| Angola | Czech Rep. | Iraq | Morocco | Singapore |
| Argentina | Denmark | Ireland | Mozambique | Slovak Rep. |
| Australia | Dominican Rep. | Israel | Namibia | Slovenia |
| Austria | Ecuador | Italy | Nepal | South Africa |
| Bangladesh | Egypt | Jamaica | Netherlands | Spain |
| Belgium | El Salvador | Japan | New Zealand | Sri Lanka |
| Belize | Estonia | Jordan | Nicaragua | Sudan |
| Benin | Eswatini | Kazakhstan | Nigeria | Sweden |
| Bolivia | Fiji | Kenya | Norway | Switzerland |
| Botswana | Finland | South Korea | Pakistan | Thailand |
| Brazil | France | Kyrgyz Rep. | Panama | Tunisia |
| Bulgaria | Gabon | Latvia | Paraguay | Turkey |
| Burkina Faso | Gambia | Lithuania | Peru | Uganda |
| Burundi | Germany | Luxembourg | Philippines | Ukraine |
| Cambodia | Ghana | Madagascar | Poland | UA Emirates |
| Cameroon | Greece | Malawi | Portugal | United Kingdom |
| Canada | Guatemala | Malaysia | Romania | United States |
| Chile | Haiti | Mali | Russia | Uruguay |
| China | Honduras | Mauritania | Rwanda | Vietnam |
| Colombia | Hungary | Mauritius | Saudi Arabia | Zambia |
| Rep. of Congo | Iceland | Mexico | Senegal | Zimbabwe |
| Costa Rica | India | | | |

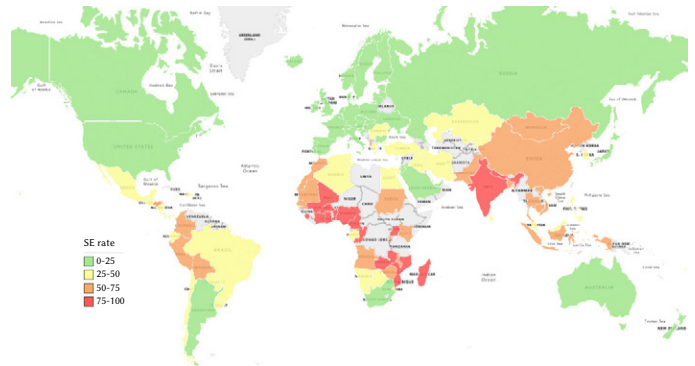


Fig. A.1. Average Self-Employment rate, 2005-2019.

TABLE A.II. VARIABLE DESCRIPTION, SOURCE AND STATISTICS

| Covariate | Code | Source | Mean | Min | Max |
|--|-------|--------------------------------|--------|--------|---------|
| Dependent | | | | | |
| Self-employed (% of total employment) | SE | ILOSTAT | 39,67 | 2,94 | 94,79 |
| GDP and components | | | | | |
| Log GDP per capita, PPP (constant 2017 \$) | GDPpc | World Bank | 9,42 | 6,62 | 11,66 |
| Agriculture, forestry, and fishing (% of GDP) | AGR | World Bank | 10,03 | 0,03 | 60,28 |
| Exports of goods and services (% of GDP) | EXP | World Bank | 41,32 | 5,32 | 228,99 |
| Imports of goods and services (% of GDP) | IMP | World Bank | 44,29 | 9,00 | 208,33 |
| Services (% of GDP) | SER | World Bank | 54,53 | 12,81 | 79,33 |
| Industry, including construction (% of GDP) | IND | World Bank | 27,18 | 0,96 | 66,76 |
| Technological Progress | | | | | |
| Patent applications, per million people | PAT | WIPO | 217,75 | 0,00 | 4212,02 |
| Individuals using the internet (% of population) | INT | World Bank | 42,18 | 0,22 | 99,15 |
| Human Capital | | | | | |
| Human Capital Index | HUC | PWT | 2,62 | 1,12 | 4,35 |
| Labor Market | | | | | |
| Labor force participation rate, female | LFF | ILOSTAT | 51,73 | 11,28 | 87,12 |
| Ratio Unemployed by Wage employees | UWE | ILOSTAT | 0,17 | 0,00 | 1,13 |
| Unemployment (% of total labor force) | UNE | ILOSTAT | 7,50 | 0,39 | 29,25 |
| Unemployment, youth (% of labor force 15-24yo) | UNY | ILOSTAT | 16,58 | 0,60 | 58,00 |
| Population | | | | | |
| Rural population (% of total population) | RUR | World Bank | 39,65 | 0,00 | 90,63 |
| Population, total in millions | POP | World Bank | 55,00 | 0,28 | 1397,72 |
| Institutions | | | | | |
| Inflation, GDP deflator | INF | World Bank | 5,66 | -26,10 | 95,41 |
| Profit tax (% of profit) | TTX | Doing Business | 44,96 | 14,10 | 285,90 |
| Labor tax and contributions (% of profit) | LTX | Doing Business | 19,17 | 0,00 | 68,00 |
| Score-Starting a business | BUS | Doing Business | 76,36 | 13,09 | 99,98 |
| Control of Corruption | COR | Worldwide Governance Indicator | 0,06 | -1,53 | 2,47 |
| Government effectiveness | GOV | Worldwide Governance Indicator | 0,15 | -2,08 | 2,44 |

ACKNOWLEDGMENT

The author wishes to thank Emilio Congregado, Jesús Crespo-Cuaresma and Concepción Román for many useful discussions. The author would like to thank the guest editor, Antonio Golpe, and two anonymous reviewers for their insightful comments that contributed substantially to the development of this paper. Financial assistance from the Spanish Ministry of Economy and Competitiveness through projects ECO2017-86402-C2-2-R, from Junta de Andalucía through Research Group SEJ-487 (Spanish Entrepreneurship Research Group – SERG) and from Research and Transfer Policy Strategy (Estrategia de Política de Investigación y Transferencia) 2020 of the University of Huelva is gratefully acknowledged.

REFERENCES

- [1] M.T. Robson, “Does stricter employment protection legislation promote self-employment?,” *Small Business Economics*, vol. 21, no. 3, pp. 309-319, 2003.
- [2] J.W. Spencer, and C. Gomez, “The relationship among national institutional structures, economic factors, and domestic entrepreneurial activity: a multicountry study,” *Journal of business research*, vol. 57, no. 10, pp. 1098-1107, 2004.
- [3] R. Torrini, “Cross-country differences in self-employment rates: The role of institutions,” *Labour Economics*, vol. 12, no. 5, pp. 661-683, 2005.
- [4] V. Kannianen, and T. Vesala, “Entrepreneurship and labor market institutions,” *Economic Modelling*, vol. 22, no. 5, pp. 828-847, 2005.
- [5] R.S. Sobel, J.R. Clark, and D.R. Lee, “Freedom, barriers to entry, entrepreneurship, and economic progress,” *The Review of Austrian Economics*, vol. 20, no. 4, pp. 221-236, 2007.
- [6] K. Nyström, “The institutions of economic freedom and entrepreneurship: evidence from panel data,” *Public Choice*, vol. 136, no. 3, pp. 269-282, 2008.
- [7] C. Bjørnskov, and N.J. Foss, “Economic freedom and entrepreneurial activity: Some cross-country evidence,” *Public Choice*, vol. 134, no. 3, pp. 307-328, 2008.
- [8] Y. Kim, W. Kim, and T. Yang, “The effect of the triple helix system and habitat on regional entrepreneurship: Empirical evidence from the US,” *Research Policy*, vol. 41, no. 1, pp. 154-166, 2012.
- [9] P. Stenholm, Z.J. Acs, and R. Wuebker, “Exploring country-level institutional arrangements on the rate and type of entrepreneurial activity,” *Journal of Business Venturing*, vol. 28, no. 1, pp. 176-193, 2013.
- [10] M.T.T. Thai, and E. Turkina, “Macro-level determinants of formal entrepreneurship versus informal entrepreneurship,” *Journal of Business Venturing*, vol. 29, no. 4, pp. 490-510, 2013.
- [11] E. Autio, and K. Fu, “Economic and political institutions and entry into formal and informal entrepreneurship,” *Asia Pacific Journal of Management*, vol. 32, no. 1, pp. 67-94, 2015.
- [12] S. Aparicio, D. Urbano, and D. Audretsch, “Institutional factors, opportunity entrepreneurship and economic growth: Panel data evidence,” *Technological forecasting and social change*, vol. 102, pp. 45-61, 2016.
- [13] M. Poschke, “Wage Employment, Unemployment and Self-Employment across countries,” *IZA DP No. 12367*, 2019.
- [14] A.F. Shapiro, and F.S. Mandelman, “Digital adoption, automation, and labor markets in developing countries,” *Journal of Development Economics*, vol. 151, pp. 102656, 2021.
- [15] C. Pietrobelli, R. Rabellotti, and M. Aquilina, “An empirical study of the determinants of self-employment in developing countries,” *Journal of International Development*, vol. 16, no. 6, pp. 803-820, 2004.
- [16] J.I. Gimenez-Nadal, M. Lafuente, J.A. Molina, and J. Velilla, “Resampling and bootstrap algorithms to assess the relevance of variables: applications to cross section entrepreneurship data,” *Empirical Economics*, vol. 56, no. 1, pp. 233-267, 2019.
- [17] K.P. Arin, V.Z. Huang, M. Minniti, A.M. Nandialath, and O.F. Reich, “Revisiting the determinants of entrepreneurship: A Bayesian approach,” *Journal of Management*, vol. 41, no. 2, pp. 607-631, 2015.
- [18] H. Low, and C. Meghir “The use of structural models in econometrics,” *Journal of Economic Perspectives*, vol. 31, no. 2, pp. 33-58, 2017.
- [19] A.E. Raftery, “Bayesian model selection in social research,” *Sociological methodology*, pp. 111-163, 1995.
- [20] C. Fernandez, E. Ley, and M.F. Steel, “Model uncertainty in cross-country growth regressions,” *Journal of applied Econometrics*, vol. 16, no. 5, pp. 563-576, 2001.
- [21] E. Ley, and M.F. Steel, “On the Effect of Prior Assumptions in Bayesian Model Averaging with Applications to Growth Regression,” *Journal of Applied Econometrics*, vol. 24, pp. 651-674, 2009.
- [22] Z.J. Acs, D.B. Audretsch, and D.S. Evans, “Why does the self-employment rate vary across countries and over time?,” Discussion Paper No. 871, *Center for Economic Policy Research*, 1994.
- [23] Z.J. Acs, S. Desai, and J. Hessels, “Entrepreneurship, economic development and institutions,” *Small Business Economics*, vol. 31, no. 3, pp. 219-234, 2008.
- [24] M. Carree, A. van Stel, R. Thurik, and S. Wennekers, “Economic development and business ownership: An analysis using data of 23 OECD countries in the period 1976–1996,” *Small Business Economics*, vol. 19, no. 3, pp. 271-290, 2002.
- [25] S. Wennekers, A. van Stel, R. Thurik, and P. Reynolds, “Nascent entrepreneurship and the level of economic development,” *Small Business Economics*, vol. 24, no. 3, pp. 293-309, 2005.
- [26] J. Crespo-Cuaresma, “How different is Africa? A Comment on Masanjala and Papageorgiou,” *Journal of Applied Econometrics*, vol. 26, pp. 1041-1047, 2011.
- [27] C. Young, “Model uncertainty in sociological research: an application to religion and economic growth,” *American Sociological Review*, vol. 74, no. 3, pp. 380-397, 2009.
- [28] E. Moral-Benito, “Determinants of economic growth: a Bayesian panel data approach,” *Review of Economics and Statistics*, vol. 94, no. 2, pp. 566-579, 2012.
- [29] E. Moral-Benito, “Model averaging in economics: An overview,” *Journal of Economic Surveys*, vol. 29, no. 1, pp. 46-75, 2015.
- [30] A. Zellner, “Bayesian estimation and prediction using asymmetric loss functions,” *Journal of the American Statistical Association*, vol. 81, no. 394, pp. 446-451, 1986.
- [31] R.E. Kass, and L. Wasserman, “A reference Bayesian test for nested hypotheses and its relationship to the Schwarz criterion,” *Journal of the American Statistical Association*, vol. 90, no. 431, pp. 928-934, 1995.
- [32] F. Liang, R. Paulo, G. Molina, M.A. Clyde, and J.O. Berger, “Mixtures of g priors for Bayesian variable selection,” *Journal of the American Statistical Association*, vol. 103, no. 481, pp. 410-423, 2008.
- [33] D.P. Foster, and E.I. George, “The risk inflation criterion for multiple regression,” *The Annals of Statistics*, pp. 1947-1975, 1994.
- [34] C. Fernandez, E. Ley, and M.F. Steel, “Benchmark priors for Bayesian model averaging,” *Journal of Econometrics*, vol. 100, no. 2, pp. 381-427, 2001.
- [35] H. Chipman, “Bayesian variable selection with related predictors,” *Canadian Journal of Statistics*, vol. 24, no.1, pp. 17-36, 1996.
- [36] E.I. George, “[Bayesian Model Averaging: A Tutorial]: Comment,” *Statistical Science*, vol. 14, no. 4, pp. 409-412, 1999.
- [37] G. Yamada, “Urban informal employment and self-employment in developing countries: theory and evidence,” *Economic development and cultural change*, vol. 44, no. 2, pp. 289-314, 1996.
- [38] S. Kuznets, “Modern economic growth,” *Yale University Press*, 1966.
- [39] S. Wennekers, A. van Stel, M. Carree, and R. Thurik, “The relationship between entrepreneurship and economic development: Is it U-shaped?,” *Now Publishers Inc.*, 2010.
- [40] E. Congregado, A. Golpe, and A. van Stel, “The ‘recession-push’ hypothesis reconsidered,” *International Entrepreneurship and Management Journal*, vol. 8, no. 3, pp. 325-342, 2012.



Ana Rodríguez-Santiago

She is a visiting researcher in the Department of Economics, Vienna University of Economics and Business. She is a PhD student in the program in Economics, Business, Finance and Computer Science (University of Huelva). Ana was granted with a FPU fellowship by the Spanish Ministry of Education. Her main areas of research are Macroeconometrics and Labor Economics and her research

focuses on Business Cycle analysis, Bayesian econometrics methods, time series analysis and finite mixture models. Her thesis focuses on the application of Bayesian methodologies on self-employment to fight model uncertainty and cluster economies based on individual characteristics.