LABOUR ECONOMICS
WORKING PAPERS

Working Paper Series
LEWP No. 3

Dissecting the impact of self-employment on unemployment: The interplay of economic performance and startup motivations

Congregado, E.
Carmona, M.
Rubino, N.

December 24

AEET
ASOCIACIÓN ESPAÑOLA
ECONOMÍA DEL TRABAJO

Dissecting the Impact of Self-Employment on Unemployment: The Interplay of Economic Performance and Startup Motivations

E. Congregado[1], M. Carmona[2], and N. Rubino*[3]

[1,2]University of Huelva and CCTH
[3]University of Tor Vergata, University of Huelva and CCTH

October 25, 2024

Abstract

This paper revisits the relationship between entrepreneurship and unemployment, challenging the prevailing notion that new startups uniformly reduce unemployment. The primary objective is to elucidate how startup motives and economic performance levels influence this relationship, emphasizing a non-linear dynamic often overlooked. Our methodological contribution lies in distinguishing between opportunity and necessity entrepreneurship, revealing that the impact on employment reduction is contingent on economic context. We find that, in economies below a certain threshold, opportunity entrepreneurs significantly contribute to lowering unemployment, while necessity entrepreneurs have the predominant effect in wealthier contexts. Although this study focuses on measurement rather than policy evaluation, it provides crucial insights for policymakers. It highlights the need for nuanced strategies that account for the motivations behind entrepreneurship and economic performance to effectively address unemployment challenges.

- **Keywords:** entrepreneurship; self-employment; unemployment; non-linear models; panel threshold models.
- **JEL:** L26; J21; J23; J24; E32; C23

*corresponding author: nicola.rubino@dege.uhu.es

1. Introduction

The popular perception among analysts and policymakers is that the creation of new start-ups (often operationalized with the new self-employed workers) contributes to reducing unemployment. Indeed, a newly self-employed individual who switches from unemployment to self-employment not only removes oneself from the ranks of the unemployed but also may create additional employment opportunities for others. However, restructuring the economy from large to small enterprises, including self-employed individuals, would not impact the overall unemployment rate as Gibrat's Law posits.

On the basis of the above arguments, promotion to self-employment is seen not only as a way of creating jobs for the new entrepreneurs themselves but also creating additional jobs in general, reducing unemployment. Thus, the effect of positive shocks in self-employment on the unemployment rate, either by turning unemployment into self-employment -using an expression popularized by Baumgartner and Caliendo (2008)- or by turning paid-employment into self-employment will depend not only on the direct effect of reducing unemployment -unemployed workers who becomes new self-employed workers- but more importantly, on the creation of additional jobs by new entrepreneurs who thus become job creators.

We cannot rule out that restructuring the economy from large to small enterprises, including self-employed individuals, would not impact the overall unemployment rate as Gibrat's Law posits. However, there is a considerable body of literature on the relative contribution of new startups compared to existing firms in terms of job creation, at the micro level. Overall, these studies appear to indicate that the newly self-employed are more important for job creation than those already self-employed (see, e.g., Acs and Armington 2002, Armington and Acs 2004, van Stel and Storey 2004, Baptista, Escaria and Madruga 2008, van Stel and Suddle 2008, and Neumark, Wall, and Zhang, 2011).

At the macro level, this idea is empirically supported by a body of empirical literature that appears to validate the existence of an inverse relationship between entrepreneurship and the unemployment rate, emphasizing that self-employment can indeed help reduce unemployment levels. This is the so-called 'entrepreneurial effect' (Thurik et al., 2008).

However, much of this work raises serious concerns.

The relationship between self-employment (or entrepreneurship) and unemployment has long been considered complex, with potential bidirectional causality. On the one hand, increased self-employment can lower unemployment by generating jobs and creating new opportunities. On the other hand, high unemployment rates can drive individuals to self-employment due to lack of job opportunities. This bidirectionality, therefore, has prompted extensive theoretical debate and empirical exploration to understand the nuances of this relationship (Faria, Cuestas and Mourelle 2010). From this perspective, it is important to pay attention to reverse causality, as its inadequate treatment can lead to biased estimates and bad inference.

A second aspect to consider relates to the possible existence of non-linearity in the relationship. Thurik et al. (2008) indicate that the relationship between self-employment and unemployment is more intense in wealthier countries, suggesting that higher income levels amplify the entrepreneurial effect on unemployment reduction. Based on certain stylized facts regarding the patterns of self-employment rates across countries, several authors have hypothesized that the different evolutions and sizes of the informal sectors may be related to a country's stage of development (see, e.g., Acs, 2006; Acs and Amoros 2008; Amoros, Cristi, Minitti 2009, Arin et al. 2015, Rodriguez-Santiago 2022). In particular, researchers often argue that countries with lower incomes face high rates of necessity-based entrepreneurial activity—self-employment as a last resort—because people are unable to find jobs. That is, people enter into self-employment because of a lack of alternatives. Extending this argument to the potential capacity of these entrepreneurs to contribute to job creation, one could argue that the predominance of this type of entrepreneur will be negatively associated with the employment intensity of self-employment.

Several studies have sought to explore the potential nonlinearity in the relationship between self-employment and unemployment. Congregado et al. (2010), Faria et al. (2010), Parker et al. (2012), and Carmona et al. (2012) each utilized diverse methodological approaches to investigate whether self-employment's effect on unemployment might fluctuate based on different economic cycles or structural conditions. For instance, Carmona et al. (2016) found that the positive or negative impact of self-employment on economic growth depends on various factors, suggesting that the relationship is indeed nonlinear and complex. More recently, O'Leary (2022) examined this relationship while distinguishing between countries with varying economic performance levels, thus introducing a new dimension to understanding potential asymmetries. His findings suggest that fostering opportunity-driven entrepreneurship in low-performing economies can enhance employment outcomes. A point of interest related to this type of literature is to understand whether certain macroeconomic conditions influence the extent to which self-employment contributes to reducing unemployment. For example, Cima et al. (2017) highlight that these processes of self-employment's impact on unemployment are more pronounced during recessions. The third critical issue in the literature pertains to the heterogeneity in self-employment. Empirical studies have traditionally employed the aggregate self-employment rate as a metric for entrepreneurship. In doing so, they overlook the significant differences that can arise from the composition of self-employment, which limits our ability to capture only a net effect.

The impact of self-employment shocks on job creation depends on several factors related to the newly self-employed. These include the type of business they start—whether they establish small and medium-sized enterprises (SMEs) or larger companies (Haltiwanger, Jarmin and Miranda 2013). Sectoral patterns also play a role, as do the proportions of employers versus own-account workers within the business landscape.[1] Building on Thurik et al. (2008), Haussen and Schlegel (2020) extend these analyses by exploring heterogeneity within solo self-employment, including differences based on gender. They use macro-level panel data from 23 OECD countries (1991–2015) and apply population-weighted vector autoregressive models and Granger-causality tests. Their findings show that solo self-employment generally has a positive indirect effect on employment in the medium to long term. However, they note that male solo self-employment tends to drive employment growth more quickly than female solo self-employment. This disparity is attributed to differences in motivation, sector selection, and business performance.

Other important source of heterogeneity refers to the economic impact of the substitution effect between paid employment and self- employment, as well as the crowding-out effect. The latter refers to the risk that new, subsidized self- employed individuals may displace or disrupt established businesses and their competition.[2]

Finally, the predominance of necessity versus opportunity entrepreneurs is another crucial factor. At the macro level, necessity entrepreneurs, who start businesses out of necessity rather than opportunity, tend to have different effects on job creation compared to opportunity entrepreneurs. Studies have shown that necessity entrepreneurs often perform less well than opportunity-driven ones.[3]

From this perspective, while opportunity-based entrepreneurs are more likely than necessity-based entrepreneurs to become employers, one could argue that positive changes in the entry rates of either type should contribute to reducing unemployment. This forms the basis of our first hypothesis: there is a negative relationship between the rate of nascent entrepreneurs and unemployment. Both nascent necessity and opportunity entrepreneurs contribute to job creation, whether by generating self-employment opportunities for themselves or by maintaining existing jobs. The impact of these self-employed individuals on unemployment reduction depends significantly on their contributions to overall employment growth. Our second hypothesis considers the potential for a non-linear relationship

[1] See, e.g., Cowling, Taylor and Mitchell 2004; Congregado, Millan and Roman 2010; Congregado, Golpe and Parker 2012; van Stel, Wennekers and Scholman 2014; and Cowling and Wooden 2021; among others.

[2] Roman, Congregado and Millan (2011).

[3] At the macro level, opportunity and necessity entrepreneurs appear to differ in their effects on job creation (Acs and Varga 2005; Wennekers et al. 2005; Wong, Ho and Autio 2005, Poschke 2013b; 2019; Fairlie and Fossen 2020; Fossen 2021). At the micro level, necessity entrepreneurs show inferior performance (see, e.g., Vivarelli 2004, or Block and Sandner 2009).

that may vary across nations based on their economic performance. Prior empirical evidence suggests that the relationship between changes in self-employment and fluctuations in unemployment rates can differ across countries and over time (Thurik et al., 2008). Indeed, the type of entrepreneurship may be country-specific (Bosma & Sternberg, 2014), meaning that local ecosystems, including the macroeconomic context, can either facilitate or impede the development of various types of startups. Thus, depending on the economic performance, the characteristics of the entrepreneurial ecosystem, and the industrial structure—specifically in terms of market concentration and specialization (Aarstad et al. 2016; O'Leary 2022)—the opportunities for entrepreneurship may be more or less fully exploited. In less competitive sectors or economies with lower economic performance, higher entry rates typically support larger-scale enterprises, which can significantly impact unemployment reduction. Conversely, in highly specialized economies with lower unemployment, new startups may primarily serve to fill existing employment gaps. In such cases, the reduction of unemployment may be influenced by transitions from unemployment; many of these transitions involve individuals with low employability, specifically categorized as necessity entrepreneurs.

The aim of this paper is to revisit the complex relationship between self-employment and unemployment, specifically focusing on how economic performance and startup motives influence this dynamic. Our estimates identify how unemployment rates are affected to a greater or lesser extent by shocks to self-employment, using the methodology of panel threshold regression models initially proposed by Hansen (1999). This model allows for an endogenous identification of the threshold at which the system switches from one regime to the other. To relax the lagged exogenous regressor assumption and the autoregressive bias, the Seo and Shin (2016) estimator are thus employed to obtain unbiased dynamic estimates of the relationship. To account for potential delays in the transition, following Gonzalez et al. (2017) we parametrize and fit a logistic transition function. Finally, we have used a slightly more flexible threshold estimator (Kremer, Bick, and Nautz (2012). This allowed us to follow a two-step instrumental procedure: adding lags of the endogenous variables as instruments for the level regressions and using the Protestant ethic variable as an instrument for the differenced series. The fact that the relationship between unemployment and self-employment is potentially bidirectional raises concerns about endogeneity bias becoming a serious issue in our results. In order to circumvent endogeneity bias, our findings undergo an additional robustness check through the use of a proxy for the influence of the Protestant ethic as an instrumental variable, uncorrelated with the error but highly correlated with entrepreneurship, inspired by Weber's thesis (Schilpzand and de Jong 2021). The contribution of this work is fourfold: i) firstly, it provides robust findings that help shed new light on the mixed evidence that previous empirical literature seemed to yield regarding the contribution of new startups to unemployment reduction; ii) secondly, it combines different approaches that configure a methodological strategy to address the analysis of these types of relationships in cross-country analysis at the macro level, without ignoring the potential non-linearity of relationships and potential endogeneity issues in a single framework; iii) thirdly, to correct the potential endogeneity bias, it employs an original instrumental variable –the penetration of the protestant ethic in each country, a variable uncorrelated with the error but highly correlated with entrepreneurship–, and iii) fourthly, Fourthly, this work recognizes the unveiling of the interplay between start-up motives and economic performance as a crucial determinant of how entrepreneurship impacts unemployment reduction. By distinguishing between opportunity-driven and necessity-driven entrepreneurship, we emphasize the nuanced dynamics that govern this relationship across different economic contexts..

Our findings reveal several key insights: i) we challenge previous literature by demonstrating that no systematic relationship exists between aggregate self-employment rates and unemployment. Instead, we propose that the relationship is non-linear and varies depending on economic context; ii) our findings also reveal that opportunity-driven entrepreneurs significantly contribute to lowering unemployment in lower-performing economies, while necessity-driven entrepreneurs dominate in higher-performing contexts.

This distinction underscores the importance of recognizing the heterogeneity of self-employment motivations. While this study centers on measurement rather than policy evaluation, the insights derived can inform policy discussions. Policymakers should promote opportunity entrepreneurship in struggling economies while reconsidering the role of necessity entrepreneurship in advanced economies to effectively address unemployment challenges.

The remainder of the paper is organized as follows. Section 2 provides an overview of the methodology used in this paper, while section 3 discusses the data used and reports our empirical findings. Finally, section 4 concludes the paper and suggests avenues for further research.

2. Empirical strategy

As mentioned, our objective is to investigate how unemployment rates respond to self-employment shocks, i.e., to explore the employment intensity of self-employment growth. In other words, we wish to investigate whether cyclical self-employment influences subsequent cyclical unemployment (and how). Examining the response of unemployment rates to self-employment shocks involves estimating the following equation:

$$\Delta u_{it} = \mu_i + \beta_1 \Delta s_{it} + \beta_2 \Delta g_{it} \tag{1}$$

where Δs, Δu and Δg are the growth rates of self-employment[4], unemployment and GDP per capita, respectively, in period t for country i.

The relationship described by equation 1 can be rewritten as a 'gap' equation, where all variables are measured in terms of the cyclical components or deviations from long-term trends. The empirical relationship can then be represented by the following equation:

$$u^c_{it} = \mu_i + \beta_1 s^c_{it} + \beta_2 g^c_{it} + \varepsilon_{it} \tag{2}$$

where $u^c_{it} = u_{it} - u^n_{it}$; $s^c_{it} = s_{it} - s^n_{it}$ and $g^c_{it} = g_{it} - g^n_{it}$. Thus, u^c_{it} captures the cyclical component of the unemployment rate, u_{it} is the current unemployment rate, and u^n_{it} is the trend level of the unemployment rate; correspondingly, s^c_{it} represents the cyclical self-employment rate (self-employment gap), s_{it} is the observed self-employment rate, and u^n_{it} is the natural self-employment rate or equilibrium rate. Finally, g^c_{it} represents the cyclical component of per capita Gross Domestic Product (or the value of the Output Gap), g_{it} stands for the GDP per capita itself, and g^n_{it} represents the (potential) output-trend.

A Hodrick-Prescott filter is used to decompose the variables in their cyclical and trend components, which are directly unobservable (Hodrick and Prescott, 1997).[5]

In our context, Equation 2 can be modified by including some limited dynamics to capture the inertia of the labor market and to avoid, in the standard fixed effect and logistic estimates the contemporaneity issue:

$$u^c_{it} = \mu_i + \beta_s s^c_{it-1} + \beta_u u^c_{it-1} + \beta_g g^c_{it-1} + \varepsilon_{it} \tag{3}$$

where the parameter for the new term u^c_{it-1} any potential long memory (but stationary) processes not controlled by the other regressors, s^c_{it-1} explains the direct cyclical contribution of cyclical self-employment to cyclical unemployment variations and ouptut gap g^c_{it-1} defines a set of possible, procyclical-dependent regressors which are neither embodied by the limited inertia of the labor markets nor by its relationship with self-employment opportunities.[6]

[4]Self-employment and entrepreneurship will be used interchangeably throughout this study. However, not all self-employment is entrepreneurial.

[5]Following Ravn and Uhlig (2002) the filtering has been executed using as a smoothing parameter the suggested $\lambda = 6.25$ for yearly frequencies, and a periodogram has been graphed post filtering to check for the consistency of the procedure.

[6]Time series analysis in panel data is subject to omitted variable bias and the issue of contemporaneity. Both

2.1. Asymmetries, endogeneity and small panel limitations

If we estimate equation 3 directly, three potential problems may arise. The first is that modeling the relationship, implicitly or explicitly, with a linear reaction function assumes, by construction, a symmetric behavior. Ignoring the existence of asymmetry when it is present may lead not only to a mis-specified model, producing bad forecasts and erroneous inferences in hypothesis testing, but also to incorrect inferences.

The second is that if self-employment depends on unemployment, there is a feedback effect from unemployment to self-employment. Thus, the linear model is not directly identifiable, i.e., there is a potential reverse causation problem. In previous empirical literature on the relationship between unemployment and self-employment, different solutions have been suggested to circumvent this problem. The most widespread solution has been the use of instrumental variables. Another proposed solution is the use of systems of simultaneous equations (see, e.g., Thurik et al., 2008). An alternative approach to this problem is suggested by Candelon, Colletaz and Hurlin (2013, p.890), who observe that the reverse causation issue can also be interpreted as a consequence of threshold effects. From this perspective, one could argue that an increase in the unemployment rate pushes people into self-employment and that self-employment then causes changes in the unemployment rate (Faria, Cuestas, and Mourelle, 2010). Following this reasoning, the use of a linear specification leads to an estimated correlation that takes into account both the influence of unemployment on self-employment in the first phase and the influence of self-employment on the unemployment rate in a second phase. The use of threshold models makes it possible to identify the influence of self-employment on unemployment in different phases (regimes).

An additional source of potential endogeneity is also caused by the contemporaneity issue. Threshold model, since their inception, would ideally take into account the weak exogeneity of the regressors as well as the potential exogeneity of the threshold variable by 'direct instrumentation' of the regressor matrix, which would be substituted by its t-periods lagged equivalent.[7] Endogeneity is also a natural consequence coming from the ideal set-up of a dynamic panel: lags of the regressand employed to frame the inertia of the employment/unemployment cycles are very likely to present parametric estimates plagued by a downward bias of the estimates.

Potential heterogeneity of the elasticity of unemployment with respect to self-employment shocks is also a relevant issue. It is often assumed that the relevant parameters are common across countries. However, we argue that the employment intensity of self-employment will differ when countries differ significantly not only in labor market institutions but also in levels of income and labor market dynamics. The introduction of asymmetries allows the estimated parameters to vary across countries so that the relationship is characterized by cross-country heterogeneity. In particular, we assume that at each date in the threshold model, the countries are divided into a small number of groups, where

would accentuate endogeneity problems affecting estimates and dispersion measures (variance). The choice for a lag in the endogenous variable is motivated by the existence of a series of historical factors which might not be captured by the covariate of the model. In a time series application furthermore, the choice of an autoregressive structure of one lag serves as an indicator of what kind of residual memory might be left into the endogenous variable after filtering and controlling for with the exogenous covariates: values of the single autoregressive parameter very close to one indicate long run memory and potential nonstationarity. On the other hand, contemporaneous values of the independent variables in the right-hand side might be correlated with the error term. This is the issue of contemporaneity. As the relevant conditions $corr(s^c_{it}, \varepsilon_{it}) = 0, \forall t$, $corr(g^c_{it}, \varepsilon_{it}) = 0, \forall t$ cannot be guaranteed, the independent variables have been introduced with a lag. This is also the most sensible choice in policy applications (like in classic time series intervention analysis) to reduce the contemporaneity issue and give a time-consistent causal interpretation to the estimates and is also obtained by construction in error correction models. Only vector auto-regressive applications with some sort of structural decomposition of the error term can address the contemporaneity issue by decomposing innovations other than introducing lags. Based on prior literature, the use of lags in any econometric specification is justified by the fact that the relationship between changes in self-employment and unemployment is not immediate. While it is true that transitions from unemployment to self-employment have a contemporaneous effect on unemployment, most of the employment created by new self-employed individuals will take time to materialize fully. Thus, the effects of changes in self-employment on unemployment may manifest with a temporal delay. This necessitates the inclusion of lags to capture both short-term and long-term effects. Previous studies often employ alternative lag structures to address this dynamic relationship (Thurik et al., 2008; Haussen and Schlegel, 2020).

[7] See, Seo and Shin (2017) for an exhaustive introduction to this issue.

each group is defined by a given elasticity of unemployment with respect to self-employment shocks, according to an observable variable.

To partially circumvent the assumption explained in the paragraph above, we resort to a first difference, generalized method of moments in the spirit of the Arellano and Bond estimator. Thorough first differencing, we would sort out of the equation any individual time invariant characteristic which would drive cross-country heterogeneity. Deeper lags of the regressand would act as an instrument for the autoregressive component, while a single lag structure would be used for the two remaining regressors. The FD-GMM estimator would thus allow for both the contemporaneous regressors and the (lagged) transition variable to be endogenous.[8]

Subsequently, a logistic function based transition application will take care of potential heterogeneity in response magnitude with respect to a gradual adjustment of the threshold with respect to an estimated central value and slope. We assume paid-employees and unemployed/inactive people need time to take a decision and actually make transition to self-employment. Although we could capture the timing of the transition, an additional degree of heterogeneity comes from considering how steep such transition could be according to the estimated threshold values detecting transitions from boom to bust phases of the cycle. Uncovering such changes in the speed of transition justifies the use of the Panel logistic threshold regression proposed by Gonzalez et al. (2017).

Finally, in the last application of this piece of research, we will tackle an alternative definition of entrepreneurship based on entry flows (the GEM Nascent entrepreneur definition). We anticipate, given some non-informative results coming from the application of the Seo and Shin (2016) estimator, that such section will exploit the small panel efficient properties of an alternative one, the Kremer, Bick and Nautz (2012) estimator.

2.2. Baseline model

Following the reasoning from the previous subsection, equation (3) should be modified to characterize a more flexible model in which parameters vary not only across individuals but also over time, allowing for the presence of asymmetries in the relationship. In our application, two competing measures are taken as potential thresholds: the level of per capita income and the unemployment rate. The heterogeneity of the relationship will then be endogenously determined in the threshold model. The model splits the sample into different groups of observations, depending on the value of the threshold variable.

A panel threshold model, in its most generic rendition and using our variables of interest, would be defined as follows:

$$u^c_{it} = \begin{cases} \mu_i + \beta_s^{(L)} s^{\mathrm{c}}_{\mathrm{it}-1} + \beta_u^{(L)} u^{\mathrm{c}}_{\mathrm{it}-1} + \beta_g^{(L)} g^{\mathrm{c}}_{\mathrm{it}-1} + \varepsilon_{\mathrm{it}} \text{ if } d_{it-1} \leq k \\ \mu_i + \beta_s^{(H)} s^{\mathrm{c}}_{\mathrm{it}-1} + \beta_u^{(H)} u^{\mathrm{c}}_{\mathrm{it}-1} + \beta_g^{(H)} g^{\mathrm{c}}_{\mathrm{it}-1} + \varepsilon_{\mathrm{it}} \text{ if } d_{it-1} > k \end{cases} \tag{4}$$

where μ_i represents the unobserved fixed effect, d_{it-1} is the lagged threshold variable, and k is the threshold parameter endogenously retrieved value. In this model, the observations are divided into two regimes, depending on whether the threshold variable d_{it-1} is smaller or greater than the threshold parameter k. The two regimes are distinguished by different regression slopes, $\beta_u^{(L,H)}$, $\beta_s^{(L,H)}$ and $\beta_g^{(L,H)}$. However, there is no reason to impose only two regimes. A more general specification, with p thresholds, could be parametrized as:

$$\mathrm{u}^c_{it} = \begin{cases} \mu_i + \beta_s^{(1)} s^{\mathrm{c}}_{\mathrm{it}-1} + \beta_u^{(1)} \mathrm{u}^{\mathrm{c}}_{\mathrm{it}-1} + \beta_g^{(1)} \mathrm{g}^{\mathrm{c}}_{\mathrm{it}-1} + \varepsilon_{\mathrm{it}} & \text{if } d_{it-1} \leq k_1 \\ \mu_i + \beta_s^{(2)} \mathrm{s}^{\mathrm{c}}_{\mathrm{it}-1} + \beta_u^{(2)} \mathrm{u}^{\mathrm{c}}_{\mathrm{it}-1} + \beta_g^{(2)} \mathrm{g}^{\mathrm{c}}_{\mathrm{it}-1} + \varepsilon_{\mathrm{it}} & \text{if } k_1 \leq d_{it-1} \leq k_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots & \\ \mu_i + \beta_s^{(p)} \mathrm{s}^{\mathrm{c}}_{\mathrm{it}-1} + \beta_u^{(p)} \mathrm{u}^{\mathrm{c}}_{\mathrm{it}-1} + \beta_g^{(p)} \mathrm{g}^{\mathrm{c}}_{\mathrm{it}-1} + \varepsilon_{\mathrm{it}} & \text{if } d_{it-1} > k_{p-1} \end{cases} \tag{5}$$

[8] See, Seo and Shin (2016) compared to Gonzalez et al. (2017).

The specification in equation (5), up to $p = 2$ will be the benchmark model used for the estimates reported in the next section. Before doing that however, we need to briefly explain the fixed effects, dynamic fixed effects and logistic smooth transition effects methods we employed to get an estimate of the $\beta_u^{(p)}$, $\beta_s^{(p)}$ and $\beta_g^{(p)}$

2.3. Hansen (1999) Fixed Effects model

Let us consider the following single-threshold model, which augments (4) with a common time invariant fixed effect μ together with the individual effect μ_i and $\mathbf{X}_{it-1} \in u_{it-1}, s_{it-1}, g_{it-1}$:

$$u_{it} = \mu + \mathbf{X}_{it-1}\left(d_{it-1} < k\right)\boldsymbol{\beta}_1 + \mathbf{X}_{it-1}\left(d_{it-1} \geq k\right)\boldsymbol{\beta}_2 + \mu_i + \varepsilon_{it} \tag{6}$$

The model belongs to a class of panel threshold models recently developed by Candelon, Colletaz and Hurlin (2013), who extended the model of Hansen (1999) to encompass the relationship between self-employment and unemployment with parameters that vary not only across individuals but also over time. This allows for asymmetries in the self-employment cyclical parameters, depending on national wealth or on labor market cyclical parameters' estimates. Equation (6) can thus be rewritten in terms of an indicator function:

$$u_{it} = \mu + \mathbf{X}_{it}\left(d_{it-1}, k\right)\boldsymbol{\beta} + \mu_i + \varepsilon_{it} \tag{7}$$

which would then define two distinct regimes:

$$\mathbf{X}_{it}\left(d_{it-1}, k\right) = \begin{cases} \mathbf{X}_{it} I\left(d_{it-1} < k\right) \\ \mathbf{X}_{it} I\left(d_{it-1} \geq k\right) \end{cases} \tag{8}$$

In substance, if k in (8) is given, than we can consistently OLS-estimate vector $\beta = (\beta_u^{(L,H)}, \beta_s^{(L,H)}, \beta_g^{(L,H)})$. When k is unknown, a grid-search needs to be done over a subset of values of the threshold variable d_{it-1}. Instead of using the whole information set, the search is thus restricted over k_{min} and k_{max}, which represent chosen quantiles of d_{it-1}. Ultimately, the chosen value of k is the one that minimizes the sum of squared residual of any possible model along the grid search:

$$\widehat{k} = \arg\min_{k} SSR(k) \tag{9}$$

As a general strategy, following Hansen (1999) once the threshold parameter is estimated, the next step is to test the null hypothesis of linearity, using the following likelihood ratio test:

$$F_1 = \frac{SSR_{\text{linear}} - SSR(\hat{k})}{\hat{\sigma}^2} \tag{10}$$

where SSR_{linear} is the sum of squares of the linear model, $SSR(\hat{k})$ is the sum of squared errors of the threshold model and $\hat{\sigma}^2$ denotes a convergent estimate of σ^2. Because under the null hypothesis, the threshold parameter $dit - 1$ is not identified, the asymptotic distribution of F_1 is not standard and does not follow a chi-squared distribution. Candelon, Colletaz and Hurlin (2013) solve this problem by using bootstrap simulations to compute critical values of the distribution of the statistics of the tests of the number of thresholds. In the cases of two or three thresholds, the same procedure is applied. If the p-value rejects the hypothesis of linearity, then we can discriminate between one and two thresholds.

Once the threshold effect is proved, the same procedure is sequentially applied to test a specification with p regimes versus $p + 1$ regimes. The process is complete when the null hypothesis is accepted.

2.4. Seo and Shin (2016) Dynamic Fixed Effects model

Once the threshold effect, the number of thresholds and its values have been determined and estimated, it is time to tackle the endogeneity issue by relaxing the assumption of lagged regressors and accounting for the downward bias in the autoregressive estimates of unemployment. As we have already mentioned, Seo and Shin (2016) offer an estimator which tackles such issues through an instrumented generalized methods of moments estimator in the first differences of the variable. Given an alternative rendition of (7):

$$u_{it} = \mu + \mathbf{X}_{it}'\boldsymbol{\beta} + (1, \mathbf{X}_{it}')\, \boldsymbol{\delta}\mathbb{I}\{d_{it} > k\} + \mu_i + \varepsilon_{it} \tag{11}$$

after first differencing, we would get:

$$\Delta u_{it} = \Delta\mathbf{X}_{it}'\boldsymbol{\beta} + (1, \Delta\mathbf{X}_{it}')\, \boldsymbol{\delta}\mathbb{I}\{d_{it} > k\} + \varepsilon_{it} \tag{12}$$

In equation (12), first differencing the model allows us to remove any idiosyncratic or common time invariant effect from the equation, allowing for an estimate of the unknown parameters β, δ and the exact value of k by GMM estimation. Most importantly, Seo and Shin (2016) prove that, under specific regularity conditions, estimates from the GMM ends up being asymptotically normal, and as such standard inference on the parameters across the difference regime applies.

As we already stated, a FD-GMM also allows to solve the contemporaneity issue both for the regressors and the threshold: this of course comes at a price. In terms of inference, testing for a threshold effect with a contemporaneous threshold values might sound an odd choice, considering any economic action, especially in the labor markets, is usually taken with some unknown degree of delay. We thus opted, to test for possible threshold behavior for both the contemporaneous values of our two candidate thresholds and their one period lagged delays.[9]

2.5. Gonzalez et al. (2017) Fixed effects logistic smooth transition model

To furtherly deepen our analysis, we ask whether or not regime switching in the unemployment/self-employment relationship takes place abruptly or whether or not it sensible to gradual changes in the transition variables. In order to do so, we employ the Gonzalez et al. (2017) panel smooth transition regression methodology (PSTR, henceforth).

Let us consider a non contemporaneous model of the kind:

$$u_{it} = \mu_i + \lambda_t + \beta_0' x_{it-1} + \beta_1' x_{it-1} f\left(d_{it-1}; \gamma, k\right) + \varepsilon_{it} \tag{13}$$

where u_{it} can be considered a scalar, x_{it-1} is a κ dimensional vector of exogenous variables, μ_1 and λ_t represent individual and time fixed effect, and ε_{it} represents the error term.

Following the authors, the transition function $f\left(d_{it-1}; \gamma, k\right)$ is a continuous function of the observable threshold variable d_{it-1} and is naturally bounded between 0 and 1. This implies that the coefficients of this two extreme values would thus be β_0 and $\beta_0 + \beta_1$. In general, the exact transition value k determines the shape of function $f\left(d_{it-1}; \gamma, k\right)$ and thus gives us an estimate of $\beta_0 + \beta_1 f\left(d_{it-1}; \gamma, k\right)$ for any individual i in any given point of time t.

Following the authors once more, the required functional form for the implicit function f could take a simple logistic shape:

[9] The linearity test however proved that the un-delayed selected threshold variables could not guarantee rejection of the null hypothesis of nonlinearity. That is why, in the remainder of the paper, all results shown consider a delayed threshold.

$$f\left(d_{it-1};\gamma,k\right)=\left(1+\exp\left(-\gamma\prod_{j=1}^{m}\left(d_{it-1}-k_j\right)\right)\right)^{-1}\quad \text{with } \gamma>0 \text{ and } c_1<c_2<\ldots<c_m \tag{14}$$

As an important standpoint, the restriction observed in (14) serve a very specific functional form: the transition needs to be monotonic and , as $\gamma\to\infty$, $f\left(d_{it-1};\gamma,k\right)$ becomes basically an indicator function, and we are back to the benchmark Hansen (1999) specification. Clearly, when $\gamma\to 0$, the model collapses to a standard panel fixed effect estimator.

As with the two preceding models, the PSTR can be generalized to serve a sequence of smooth transitions to different states. In that case, the model transitions additively:

$$u_{it}=\mu_i+\lambda_t+\beta_0'x_{it-1}+\sum_{j=1}^{r}\beta_j'x_{it-1}f_j\left(d_{it-1}^{(j)};\gamma_j,k_j\right)+\varepsilon_{it} \tag{15}$$

in (15), the model additively transition across r regimes. Again, if $\gamma\to\infty$ and $d_{it-1}^{(j)}=d_{it-1}$, we are back to the Hansen discrete threshold modelling with $r+1$ transitions/regimes.

The opportunity to make use of this PSTR modelling comes with the possibility to run a threshold starting from an initial parameter value for k, a positive, nonzero value for g and then recursively optimizing the fit of the logistic by updating new values for k. Furthermore, following the authors we offer an alternative to the bootstrapped, residual based Hansen nonlinearity tests we discussed two section before to introduce an alternative, nonlinear Lagrangian Multiplier test which is based on a Taylor expansion around a zero value for the slope of the logistic, essentially testing $\mathrm{H}_0:\gamma=0$.

2.6. The Kremer, Bick and Nautz (2012) estimator

The Kremer, Bick and Nautz (2012) dynamic panel threshold estimator builds on Hansen (1999) in a way close to Seo and Shin (2016), but presents the possibility of performing the analysis through a 2SLS estimator (Caner and Hansen 2004) which allows us to qualify our protestants variable as an instrument for the first step difference equation, and the lag of the objective variable as an instrument for the level equation. This allowed us to partially circumvent the limitations on the trimming value which depend on the small dimensions of our panel, at the cost of resorting to a two step analysis instead of a single GMM estimate.

The Kremer, Bick and Nautz (2012) estimator presents an additional advantage over the Seo and Shin (2016). Instead of first differencing or demeaning this estimator controls country-specific fixed effects by subtracting an average of all future values of a variable (say y for simplicity):

$$y_{it}^{*}=\sqrt{\frac{T-t}{T-t+1}}\left[y_{it}-\frac{1}{T-t}\left(y_{i(t+1)}+\ldots+y_{iT}\right)\right] \tag{16}$$

error terms will thus be uncorrelated and we will be sure the distributional assumptions from Hansen (1999) were not violated. This comes with a small disadvantage: the model targets a single regime-dependent variable, mantaining the linearity assumption for the rest of the model. Besides the major differences seen above, the Kremer, Bick and Nautz (2012) estimator is based on a canonical threshold model regression:

$$y_{it}=\mu_i+\beta_1'z_{it}I\left(q_{it}\leq\gamma\right)+\beta_2'z_{it}I\left(q_{it}>\gamma\right)+\varepsilon_{it}, \tag{17}$$

where, subscripts $i=1,\ldots,N$ represent the country and $t=1,\ldots,T$ index the time. μ_i is the country-specific fixed effect, and the error term is $\varepsilon_{it}\overset{\text{iid}}{\sim}\left(0,\sigma^2\right).I(\cdot)$ is the categorical indicator that defines the regime switch conditional on the values of threshold variable q_{it}, and the threshold level $\gamma \cdot z_{it}$ is a m-dimensional vector of explanatory regressors which can include lagged values of y and other endogenous variables.

Together with Equation (17), the model requires a set of instrumental variables z_{1it} and after the initial differentiation can either be estimated by GMM or with a two stage least square procedure (2SLS), which permits a discretional selection of the instrumental variables required for the first and the second step of the computation. As in Seo and Shin (2017), the value γ is computed recursively and being the model slightly simplified in terms of its regime-dependent components (only n_{it} is allowed to be regime dependent) we managed to increase the trimming value up to the canonical and literature-based 10% without any further issue.

3. Data and Results

In this section, after a brief description of our data, the empirical results are presented in several steps. First, we discuss the stationary properties of the self-employment/entrepreneurship and unemployment series. Second, we check the null of linearity and in the case of rejection, seek the 'best' threshold variable. Third, we report estimates for the static (Hansen, 1999) dynamic (Seo and Shin, 2016) and smooth (Gonzalez et al., 2017) transition models of the relationship for the different regimes defined by the selected threshold variable. Fourth, we utilize a different metric for entrepreneurship, specifically nascent entrepreneurs, measured using data extracted from the Global Entrepreneurship Monitor, with the aim of assessing whether the significance of the relationship may be influenced by this factor. Finally, we employ the same dataset to explore the heterogeneity of self-employment by taking into account the start-up motive. In both cases, the characteristics of the panel lead us to use the Kremer, Bick, and Nautz (2012) estimator.

3.1. Data

As noted above, our goal is to examine whether positive entrepreneurship shocks can cause reductions in unemployment and whether the employment intensity of these shocks is asymmetric across countries, depending on economic conditions and the predominant type of start-up. To this end, we use a sample of over three decades of annual time-series data, covering the 1990-2020 period, for 23 OECD countries, including Australia, Canada, Iceland, Japan, New Zealand, Norway, Switzerland, the EU-15 and the United States. As noted, and as in most previous studies, in this article, entrepreneurship is operationalized as self-employment, reflecting data availability at the time-series level (Parker, 2009).[10] Self-employment series come from two different data sources. Self-employment rates are taken from World Bank database, where self-employment rates are defined as a percentage over total employment. Unemployment rate is also sourced from the World Bank databank website and represented as a percentage over the total labor force. The second proxy for self-employment is represented by the GEM TEA rate taken from the Global Entrepreneurship Monitor. The opportunity and necessity entrepreneurial series were collected from the same source. GDP per capita is taken from the World Bank database and is measured in thousands of US dollars at 2017 constant prices (constant PPPs). The last time series included in our analyses is the harmonized unemployment rate, which is provided by OECD Main Economic Indicators. Finally as previously mentioned, we employ the series on protestantism as instrumental variable in the relationship. This time series were sourced from www.theglobaleconomy.com. The countries for whom we could retrieve the series for instrumentation

[10] In this respect, we are aware that entrepreneurship is a multifaceted concept and self-employment rates or business ownership rates are limited as proxies (Wennekers and Thurik, 1999; Congregado, 2008; Iversen, Jorgensen and Malchow-Moller, 2008). Although self-employment or Total Entrepreneurship Activity (TEA, henceforth) may be close to the Knightian entrepreneur, these are far from being optimal indicators of other vectors that define the entrepreneurial function. In any case, self-employment rates and the TEA are the only variables available to measure entrepreneurship at the country level over long time periods.

were: Australia, Austria, Germany, Denmark, Spain, Finland, France, Great Britain, Ireland, Netherlands, Norway, New Zealand, Portugal, Sweden, United States. Given the much smaller magnitude of the database and to avoid collinearity resulting in null results the objective variable had to be switched from cyclical to complete unemployment rate.

As our interest lies in analyzing transitions related to the downturns and upturns of the business cycle, all the variables have been filtered in order to disentangle cyclical fluctuations from any sort of trend may it be deterministic or stochastic, following Hodrick and Prescott (1997) and Ravn and Uhlig (2002). The selected bandwidth which would allow us to disentangle fluctuations from trends, given our yearly frequency was $\lambda = 6.25$.

We offer a brief look at the periodograms for the cyclical components of self-employment, unemployment and GDP per capita in Australia in Figures 1, 2 and 3. Although with different degree of success, reversion to the original value of -6.00 might be an issue, especially at higher frequencies, to the right of the second vertical line. One possible explanation is of course given by the nature of the procedure, since the Hodrick-Prescott is a high pass filter. Furthermore, we just not need to consider every single country as an isolated study case, but need to consider overall estimates once time invariant idiosyncracies and common effects will be taken out of the equation by our estimators. To circumvent the issue of establishing a specific, country by country bandwidth-length to increase the accuracy of the filter, in the next session we will present the panel unit root results for the cyclical component of each variable. This will give us an idea of the overall stationarity of the cyclical component panel, and give us weak confirmation of the validity of the average filtering we executed by considering a panel-invariant bandwidth λ, identical for each of the 23 countries included in the analysis. To get some preliminary understanding of the strength of the relationship between cyclical unemployment and cyclical self-employment and to determine how heterogeneous such relationship is across cross sectional units, Table 1 reports the inter-country correlations between the two measures.

Figure 1: Periodogram, cyclical self-employment in Australia

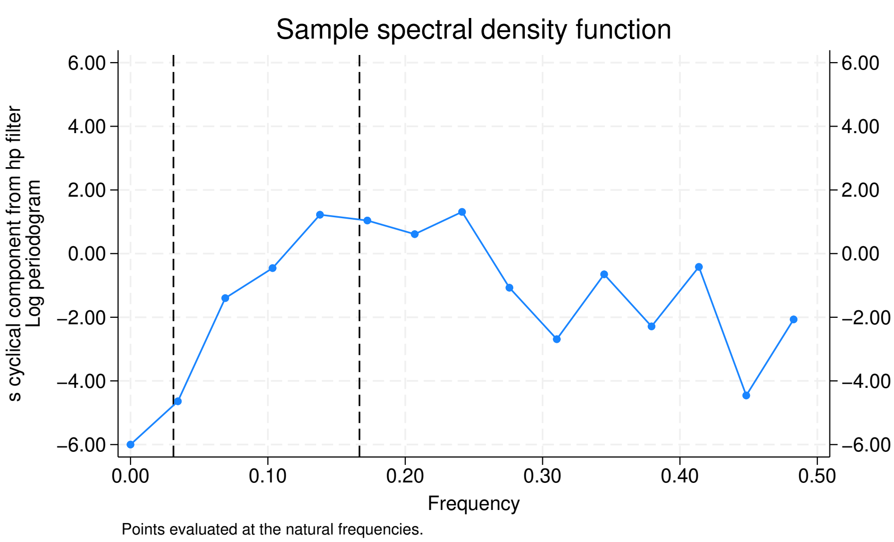


Figure 2: Periodogram, cyclical unemployment in Australia

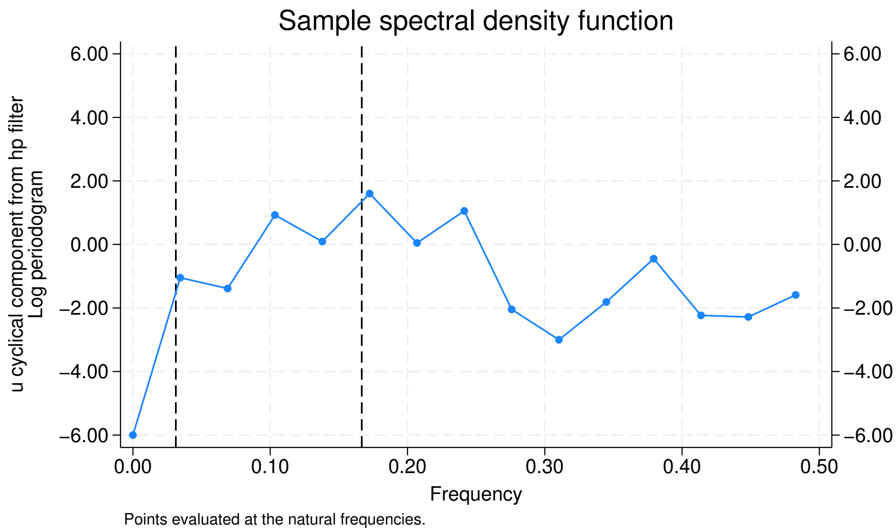


Figure 3: Periodogram, cyclical GDP per capita in Australia

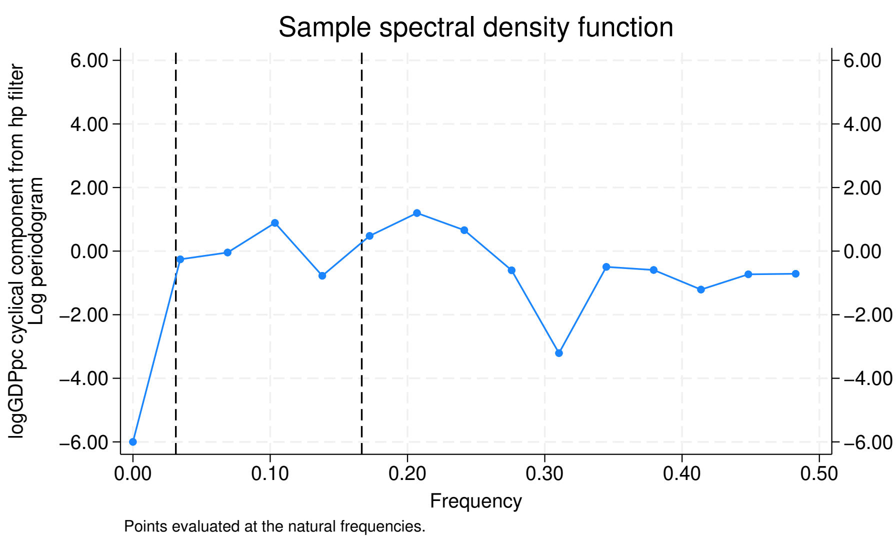


Table 1: Correlations between cyclical unemployment and cyclical self-employment

Country (I)	ρ (II)
Australia	0.420
Austria	-0.07
Belgium	0.586
Canada	0.209
Switzerland	0.155
Germany	0.614
Denmark	0.405
Spain	0.224
Finland	0.404
France	0.435
Great Britain	-0.242
Greece	0.371
Ireland	0.345
Iceland	-0.260
Italy	0.085
Japan	-0.029
Luxembourg	0.220
Netherlands	0.582
Norway	0.011
New Zealand	0.067
Portugal	0.420
Sweden	0.424
United States	0.143

The table reports the raw, pair-wise correlation between the filtered cyclical self-employment and the cyclical unemployment variables.

3.2. Unit root analysis

A crucial point of our estimation methodology has to do with the need to work with stationary cyclical variables, since the estimators we shall employ later in the article where normally thought with a fixed T and $N \to \infty$ in mind.[11]

For this purpose, we use a battery three panel unit root tests: the heterogeneous stationarity test proposed by Hadri (2000), the Homogeneous hypothesis unit root test proposed by Levin, Lin and Chu (2002) and finally heterogeneous hypothesis unit root test by Im, Pesaran and Shin (2003). Since we are already working with the cyclical components of the original filtered variables, we thought about omitting additional testing for higher order of differentiation, such as $H_0 : {u_t}^c \sim I(2)$ vs $H_1 : {u_t}^c \sim I(1,0)$, and focus on the the $I(1)$ vs $I(0)$ case. However, given the heterogeneous efficiency of the first-step Hodrick-Prescott filtering, we opted for a complete anaylsis of the series up to two orders of differentiation, in order to get some conclusive inference on the order of integration of the panel.

The outcome of the tests is reported in Table 2. Both the Levin, Lin and Chu (2002) and the Im, Pesaran and Shin (2003) unit root tests allowed us to reject both the Homogeneous and Heterogeneous null hypotheses on the order of integration of the process up to $I(2)$, underlying the stationarity of the three series. The Hadri Stationairty test, even when we imposed a relatively long Bartlett kernell lag window (around 9 lags) was also unable to reject the null of non-stationarity at 1%.

All in all, we are able to comfortably consider the overall panel stationary. As a caveat, all the testing equations were fitted with a non-zero constant as a unique deterministic component. That comes from the fact that filtering the series naturally de-trends them. Additional results on the order of integration of the data are, however, available on request.

[11] Although filtered series would be deemed as stationary by definition, the statistical qualities of the stationary component depend on the filtration parameters. For such reason we opted to run a preliminary analysis to test for the stationarity of the filtered series.

Table 2: Panel integration analysis. Variables: cyclical self-employment rate, cyclycal unemployment rate and cyclical GDP per capita. Panel obs.: N = 23; T = (1991-2019).

Series (I)	Test (II)	H_0 (III)	H_1 (IV)	Lags (V)	Stat (VI)	Decision (VII)
${s_t}^c$	ΔLLC	Hom I(2)	Hom I(1 or 0)	0	-22.756***	not I(2)
	LLC	Hom I(1)	Hom I(0)	6	-24.524***	Hom I(0)
	ΔIPS	Hom I(2)	Het I(1 or 0)	0	-22.194***	not I(2)
	IPS	Hom I(1)	Het I(0)	6	-18.552***	Het I(0)
	ΔHadri	Hom I(1 or 0)	Het I(2)	0	-4.525	not Hom I(1 or 0)
	Hadri	Hom I(0)	Het I(1)	0	-4.154	Hom I(0)
${u_t}^c$	ΔLLC	Hom I(2)	Hom I(1 or 0)	0	-25.413***	not I(2)
	LLC	Hom I(1)	Hom I(0)	0	-19.371***	Hom I(0)
	ΔIPS	Hom I(2)	Het I(1 or 0)	0	-19.470***	not I(2)
	IPS	Hom I(1)	Het I(0)	0	-15.616***	Het I(0)
	ΔHadri	Hom I(1 or 0)	Het I(2)	0	2.971	not Hom I(1 or 0)
	Hadri	Hom I(0)	Het I(1)	0	2.313	Hom I(0)
${g_t}^c$	ΔLLC	Hom I(2)	Hom I(1 or 0)	0	-28.111***	not I(2)
	LLC	Hom I(1)	Hom I(0)	0	-24.463***	Hom I(0)
	ΔIPS	Hom I(2)	Het I(1 or 0)	0	-5.590***	not I(2)
	IPS	Hom I(1)	Het I(0)	0	-3.752***	Hom I(0)
	ΔHadri	Hom I(1 or 0)	Het I(2)	0	2.284	not I(2)
	Hadri	Hom I(0)	Het I(1)	0	1.661	Hom I(0)

*** Null rejection at 1% level. All test equations are augmented with their cross-sectional averages and include a nonzero intercept as the unique deterministic component. (I) Tested time series; (II) Performed tests: Levin, Lin and Chu -LLC-, Im, Pesaran and Shin -IPS- and Hadri, run in first differences -Δ- and, where appropriate, levels to follow the Dickey and Pantula (1987) suggestion; (III-IV) Null and alternative hypothesis. *Hom* indicates homogeneity of the hypothesis (the hypothesis entails all cross sections), *Het* indicates heterogeneity of the hypothesis: the associated order of integration is valid for at least one cross section of the panel; (V) Most parsimonious lag structure suggested by the AIC, BIC and HQIC criteria; (VI-VII) Test statistic and decision under the null.

3.3. Linearity test, threshold selection and Hansen (1999) Fixed Effects Estimates

Once the cyclical components of the time series have been extracted and the overall stationarity of the panel checked out, we must check the null of linearity and determine the 'best' threshold variable.[12] We consider two candidates in turn: the level of income per capita and the unemployment rate, $d_{it-1} = g_{i,t-1}, u_{i,t-1}$.

On the one hand, it appears logical that past unemployment rates should play a role in regime switching: a high unemployment rate is a symptom of a low job offer arrival rate, lower reservation wages, and in general, a higher propensity of the unemployed to react to positive incentives to undertake self-employment. On the other hand, it is also possible that the self-employed sector is larger in countries that have lower GDP per capita. If we take entrepreneurship and self-employment to be equivalent, high levels of self-employment may result from low salaried employment opportunities (see, e.g., Blanchflower (2004), Acs (2006), Shane (2009), Congregado, Golpe and Carmona (2010) or Loayza and Rigolini (2011), among others).[13] The conceptual distinction between necessity and opportunity entrepreneurs is also closely related to this view.[14] On the basis of the above considerations, one could argue that the employment intensity of self-employment should differ across countries, depending on wealth. Therefore, the threshold variable can also be defined as the level of income per capita, according to the arguments exposed above.

As usual in the estimation of panel threshold regression models, we discriminate between these two candidates according to a statistical criterion. In particular, we will choose the variable that i) minimizes the sum of squared residuals (Hansen, 1999) and ii) leads to the strongest rejection of the linearity hypothesis, based on the results of the already mentioned F_1 statistic. There is a sensibility issue: the grid-search necessary to calculate the transition value depends on the width of the trimming window. In order to accommodate such possibility, we test for all possible trimming values in the neighborhood between a $5th$ percentile to $95th$ percentile rolling window, reducing it slowly until the rolling window equals a total number of observations as those comprised between the $25th$ and the $75th$ percentile. The results of the grid-search and the corresponding values for the linearity test, the optimal threshold value k together with the residuals sum of squares of each bootstrapped model and its corresponding probability value are reported in Table 3.

The likelihood ratio test F_1 clearly leads to the rejection at 5% of the null hypothesis of linearity of the relationship across all possible trimming choices but only when the threshold variable is income per capita. Some weak hint at threshold behavior could be find, close to five percent, only when the selected threshold for lagged unemployment was estimated at around a 7.03% unemployment rate (close to the overall average of the panel) with a 0.2 percentiles regression window. This result suggests estimating the model in a non-linear form, using national economic performance as the threshold variable and excluding the unemployment rate as a candidate. According to Hansen's procedure, it is necessary to estimate and test two thresholds, and so on, until the corresponding F-test is statistically non-significant.

Following this strategy, we re-ran the test for a trimming window of 0.1, 0.2 and 0.3 percentiles, reducing the grid-search amplitude to save some computational time and avoid computational errors given by the reduced number of total observations at our disposal. The null of a single threshold against an alternative three-regimes specification was never rejected under any canonical confidence level.

Based on its significance level, we thus checked for the only optimal threshold. The threshold

[12] According to Candelon, Colletaz and Hurlin (2013) the only constraint is that the threshold variable cannot be time-invariant. We also add, naturally, that in order to control for the issue of contemporaneity-induced endogeneity the all the threshold variables have been lagged once.

[13] See Gërxhani (2004) for a survey of informality in developed and less developed countries.

[14] Acs (2006) suggests the use of the ratio of opportunity to necessity entrepreneurship as a useful indicator of the relationship between entrepreneurship and economic development.

results, their confidence intervals and the p-values of the corresponding linearity tests are reported in Table 4. The threshold estimate indicates when the transition between the two regimes occurs. For example, if real GDP per capita of a country exceeds the threshold (37898$), the country in question switches from the lower to the upper regime. By contrast, the first regime is obtained when national per capita income is below 37898$. As we tabulate the outcomes of the model, 31% of the observations appear to belong to the lower regime, around 69% to the upper one. However, countries that switch from one regime to the other combine episodes below and above the threshold. That is, any country may experience episodes where the per capita income is either above or below the threshold value.

Table 3: Hansen (1999) bootstrapped Linearity tests: Linearity against a single threshold

Trimming (I)	Statistic (II)	g_{it-1} (III)	u_{it-1} (IV)
0.1	k	30931$	7.030%
	RSS	198.057	201.086
	F_1	23.710	14.070
	$P-value$	0.039**	0.1392
0.2	k	37898$	7.030%
	RSS	198.902	201.086
	F_1	20.990	14.070
	$P-value$	0.050**	0.059*
0.3	k	37898$	4.970%
	RSS	198.902	204.730
	F_1	20.99	2,860
	$P-value$	0.017**	0.707
0.4	k	37898$	4.330%
	RSS	198.902	203.252
	F_1	20.99	7.360
	$P-value$	0.020**	0.265
0.49	k	30931$	2.950%
	RSS	198.057	201.851
	F_1	23.71	11.690
	$P-value$	0.027**	0.221

** Null rejection at 5% level. All tests have been bootstrapped with 5000 repetitions. The null tested is of linearity against the null of a nonlinear, jump-like threshold transition. Column (I): trimming value as a proportion of the whole series; Column (II): revelant statistics; Column (III) lagged GDP per capita as a threshold; Column (IV) lagged unemployment rate as a threshold.

Table 4: Threshold selection

Trimming (I)	k (II)	$CI-$ (III)	$CI+$ (IV)	P-values (V)
0.1	30931$	30916$	30966$	0.039
0.2	37898$	37550$	37918$	0.050
0.3	37898$*	37853$	37918$	0.017*
0.4	37898$	37550$	37918$	0.020
0.49	30931$	30595$	30966$	0.027

* Selected threshold. Column (I): trimming value as a proportion of the whole series; Column (II): threshold value; Column (III) upper confidence interval; Column (IV) lower confidence interval.

Let us now observe the result of the Panel Threshold application, which correspond to the threshold value $k = 37898\$$ with a trimming window of 0.3. Considering the residual some of squares of the F_1 for $d_{it-1} = g_{it-1}$ and conditional on the rejection of the null hypothesis, another threshold value would be identifiable, with $k = 30931$ at the 0.1 trimming value.[15] We thus present the results for both threshold values for the lagged GDP per capita threshold. In order to keep the analysis cohesive, we also offer present the estimates for the only (marginally) statistically significant nonlinear model which comprises the lagged unemployment rate as a threshold, when its value is equal to $k = 7.03\%$.

Results of the panel threshold estimates can be seen in Table 5.

Our results do not confirm the existence of a systematic (and inverse) relationship between new (aggregated) entries into self-employment and the unemployment rate, as a significant portion of previous literature has pointed out (Thurik et al., 2008; Apergis and Payne, 2015; Haussen and Schlegel, 2020; Komninos et al., 2024). Instead, our results indicate that the relationship is not linear and thus does not operate under certain regimes, in line with the results of O'Leary (2022). Specifically, there is statistically significant evidence of a contribution from new entrepreneurs to unemployment reduction, but only in economies classified under the high economic performance regime (defined as having a per capita income level above the threshold).

Furthermore, estimates also show the existence of inertia in unemployment, such as dependence of unemployment on prior labor market dynamics. This effect however appears stronger, both in columns (II) and (III) of Table 5, in countries when their income is above the endogenously retrieved GDP per capita threshold threshold.[16]

There is however a strong caveat to direct-effect or even causal interpretations that we will tackle in the next subsection: the built in endogeneity given by the presence of an autoregressive component (whose magnitude might probably be upward-biased) and the assumed exogeneity of the regressors and the threshold value.

[15] We already point out that this might be a sign of smooth transitional behavior: although the double threshold F_2 test null hypothesis of double threshold behavior was rejected, the existence of more than one feasible threshold conflicting with Hansen (1999) selection choice might be indicative of a gradual transition from a lower to an upper state. We shall elaborate and present some evidence of such transition in the section dedicated to the Gonzalez et al. (2005-2017) estimator

[16] At this point, we might be tempted interpret the previous puzzling findings on unemployment persistence as follows: when cyclical unemployment is very high, it usually has a certain persistence, becoming itself the source of further persistence. As a matter of fact, this is also visible in Column (IV) of Table 5: if the unemployment rate is used as a threshold, the lagged value of the endogenous variable used as a regressor does jump up in value and is perhaps the highest in magnitude across the three specifications.

Table 5: Hansen (1999) bootstrapped threshold fixed effect estimates

$Thresholds$ (I)	g_{it-1} (II)	g_{it-1} (III)	u_{it-1} (IV)
Trimming	0.3	0.1	0.2
k	37898$	30931$	7.03%
${u_{it-1}}^{cycleL}$	0.175***	0.010	0.068
	‘(0.069)	(0.115)	(0.082)
${s_{it-1}}^{cycleL}$	-0.116	-0.002	0.179*
	(0.131)	(0.193)	(0.101)
${g_{it-1}}^{cycleL}$	-29.825***	-46.094***	-15.601***
	(4.424)	(7.535)	(3.132)
${u_{it-1}}^{cycleH}$	0.302***	0.287***	0.374***
	(0.061)	(0.050)	(0.055)
${s_{it-1}}^{cycleH}$	-0.231***	-0.203***	-0.132
	(0.102)	(0.089)	(0.133)
${g_{it-1}}^{cycleH}$	-9.239***	-11.185***	-14.216***
	(2.441)	(2.200)	(2.995)
T	28	28	28
N	23	23	23
R^2	0.280	0.283	0.270

*** Null rejection at 1% level; ** Null rejection at 5% level; * Null rejection at 10% level. Country fixed-effects included. Column (I): relevant statistics; Column (2),(3), (4): estimated parameters for the indicated threshold variable value. The apex values *cycleL* and *cycleH* indicate that the cyclical regressor component belongs to either the regime below (L) or above (H) the threshold. Standard Errors in parentheses. Objective variable: cyclical unemployment rate, ${u_{it}}^c$.

3.4. Seo and Shin (2016) dynamic FD-GMM estimator

To deal with the endogeneity issues deriving from the presence of an endogenous dynamic regressor and the contemporaneity issue of the threshold variable, this section presents the estimates of the First Differenced, Generalized Methods of Moments based estimator from Seo and Shin (2016). To get a more precise idea of the severity of the structural endogeneity issue in the previous section, we will allow deeper lags of the endogenous variable to act as instrument for the estimates after first differencing and allow the contemporaneous value of GDP per capita and unemployment to act as thresholds.[17]

The results for the Dynamic Panel estimation from Seo and Shin (2016) are reported in Table 6. As before, the threshold value, conditional on a set of trimming intervals, was selected according to its statistical significance. Column (I) in this regard, differs from column three because, once a contemporaneous GDP per capita variable is selected as a threshold, the optimal threshold value, 47880$, stands closer than ever to the panel average, equal to 45984$. This allows for the majority of the observations to switch over to the lower regime (58% against the 30% from the Hansen (1999) estimates) and effectively reduces the inertia effect of unemployment. The sum of all periods spent under the threshold ends up weighting more than the sum of all periods spent over the threshold among all 23 nations our database considers. If we accept contemporaneity of the threshold, the effect is still present in the upper regime, but its anti-cyclical impact on the present value of the unemployment cycle is now less than half the pro-cyclical impact of self-employment in high regime countries.

Once again, estimates of the impact of cyclical self-employment on cyclical unemployment are statistically not distinguishable from 0 in the low GDP per capita regime. As before, we verify again that there is no systematic and inverse relationship between new entries into self-employment (in aggregate terms) and the reduction of the unemployment rate; rather, this relationship is only statistically significant when economies are in the high economic performance regime (i.e. a nonlinear relationship).

All in all, consistency of the upper regime self employment leads us to conclude that, as it might be expected, positive shocks to self-employment put downward pressure on unemployment not only because some unemployed individuals create their own jobs (self-employed workers) but because some of them become job creators, taking advantage of such possibility in the best conditions possible, offered by the higher regime.

The result remains the same, such that there is only evidence of a systematic relationship between new entries into self-employment in situations where the economy is positioned in period t within the high economic performance regime.

To keep the analysis consistent, we present, in Column(III), the best possible dynamic representation for the most statistically representative unemployment threshold. Besides the threshold value 4.033% being much lower relative to the average of the panel (7.349%) the estimates did no catch any meaningful correlation, aside from some weak persistence of the endogenous variable in the upper regime.

[17] We might allow an alternative to direct usage of lagged regressors in the models, allowing for contemporaneity of both the threshold variable and the cycle components, but in terms of dynamics of the labor market, given that changes in the business cycle might take some time to reflect into actual values, we will leave the cyclical regression in their lagged form.

Table 6: Seo and Shin (2016) FD-GMM dynamic estimator

$Thresholds$	g_{it-1}	g_{it}	u_{it-1}
(I)	(II)	(III)	(IV)
Trimming	0.02	0.02	0.03
k	54587$***	47880$***	4.033%***
	(5096)	(5435)	(1.219)
${u_{it-1}}^{cycleL}$	-0.156	-0.249	-2.826
	(0.504)	(0.254)	(1.877)
${s_{it-1}}^{cycleL}$	0.287	0.366	0.125
	(0.703)	(0.814)	(0.909)
${g_{it-1}}^{cycleL}$	-38.502***	-22.839*	23.387
	(14.543)	(12.551)	(32.626)
${u_{it-1}}^{cycleH}$	-1.488	1.980***	3.296*
	(1.669)	(0.541)	(1.891)
${s_{it-1}}^{cycleH}$	-6.314***	-5.529***	-1.252
	(2.748)	(1.620)	(1.014)
${g_{it-1}}^{cycleH}$	-37.811	24.075	-37.508
	(47.451)	(26.594)	(37.866)
T	28	28	28
N	23	23	23

*** Null rejection at 1% level; ** Null rejection at 5% level; * Null rejection at 10% level. Country fixed-effects included. Column (I): relevant statistics; Column (2),(3), (4): estimated parameters for the indicated threshold variable value. The apex values *cycleL* and *cycleH* indicate that the cyclical regressor component belongs to either the regime below (L) or above (H) the threshold. Standard Errors in parentheses. Objective variable: cyclical unemployment rate, ${u_{it}}^c$.

3.5. Seo and Shin (2016) FD-GMM estimator: tackling reverse causality with additional instrumentation

Results obtained in the previous section underline the relative importance of upper-regime self employment rates on job creation once a given GDP threshold has been passed. A direct inference in terms of conditional elasticity however, appears to be difficult as the estimated self-employment coefficient overlaps the results of the Hansen Panel estimates by a tenfold. This upward bias in visible in Table 6, where a 1% increase in cyclical self employment in period t in the upper GDP regime would correspond to more or less a 6% decrease in cyclical unemployment (with an elasticity of -6.314^{***} for the cyclical lagged self-employment coefficient).

In order to account for possible additional endogeneity issues, possibly driven by omitted variable bias and depending on the quality of the best descriptors for unemployment rates, this section exploits the GMM methodology of Seo and Shin (2016) to add an additional instrument in the differencing procedure of the estimator, which we would assume to be correlated enough with our variable of interest, self employment.

The selected instrument was thus the percentage of Christian Protestants in all available countries and time periods.

The relationship between entrepreneurship/self-employment and unemployment is influenced by reverse causality, where both variables can affect each other. While entrepreneurship is often viewed as a driver of job creation (Audretsch et al., 2001), high unemployment can motivate individuals to pursue self-employment, especially during economic downturns (Parker, 2009). This bidirectional relationship complicates the analysis, as rising unemployment may lead to increased self-employment rates (Thurik et al., 2008). Failing to address reverse causality can result in bad inference and biased estimates, risking overestimating or underestimating the effects of entrepreneurship on unemployment.

To accurately estimate the relationship, using instrumental variables (IV) is essential for addressing endogeneity issues. A suitable IV should be strongly correlated with entrepreneurship but exogenous to unemployment. Protestantism serves as an effective instrument, as Weber (1905) linked it to higher entrepreneurial activity through cultural norms emphasizing individual responsibility and discipline. Empirical studies have supported this correlation, indicating that Protestantism does not directly correlate with unemployment rates, thus satisfying the exogeneity condition for IVs (Becker & Woessmann, 2009; Guiso et al., 2006).

The Protestant ethic encourages traits conducive to opportunity-based entrepreneurship, distinguishing it from necessity entrepreneurship, which arises from economic need (Williams & Nadin, 2010). While Protestant values foster a proactive entrepreneurial spirit, they are less associated with addressing unemployment directly, focusing instead on promoting economic agency and growth (Audretsch & Thurik, 2001). Consequently, the connection between Protestantism and entrepreneurship is more closely related to opportunity-driven initiatives rather than necessity entrepreneurship, underscoring the importance of considering these dynamics in economic analyses.

In sum, by incorporating Protestant Ethic Penetration, we mitigate endogeneity issues by providing a more exogenous source of variation in entrepreneurial activity. This approach is especially valuable in studies where measuring the impact of entrepreneurship is challenging due to potential reverse causality or confounding factors.

After checking for pairwise correlation and statistical significance at 1% after the Bonferroni Correction, the calculated, significant value of the correlation between s_t and p_t (the percentage of Christian Protestant in any given year for all available subsetted countries) amounted to -0.5686^{*}. This section reports the estimates of the adjusted, Seo and Shin (2016) FD-GMM dynamic estimator with the above-cited additional instrument paired with some deeper lags of the dependent variables.

After some remodeling of the data, the number of observations at our disposal results in a panel of 330 entries, distributed along 22 periods, from 1992 to 2013, across 15 different countries.

To ensure adherence with the results from the previous section and show the non-marginal effects on the magnitude of the coefficients after instrumentation, Table 7 reports the estimated values for the best fitting model with and without instrumentation, considering the lag of GDP as a threshold once more. Comparing the results of Table 7 for the instrumented version and the non-instrumented model, we observe how the upward bias in the magnitude of the coefficient for ${s_{it-1}}^{cycleH}$ now shows the expected sign and a more reasonable, literature-observed magnitude, as a 1% increase in self-employment at time t will now suggest a 0.6% decrease in unemployment. The results for self-employment in the instrumented estimates stay consistent for all possible trimming windows, conditional on the amount of available observations left in the database after balancing the panel to run the estimates properly, with a threshold value consistently estimated in the neighborhood of 48000$.[18]

Together with the effect of the lagged regressors and the use of an FD-GMM estimator, the use of additional instrumentation further reduces the endogeneity bias up to the highest degree possible in an application of our kind. This result, as we will see in the long-run multiplier section, stands slightly above the naive linear panel applications but is dwarfed by the results of the nonlinear application, that we shall present in the next section.

In the short term the -0.553 estimate for the short run elasticity relative to cyclical self employment co-movements with unemployment in Table 8 shows an effect of job creation which almost doubles the -0.331 computed by the Standard Hansen Estimates and visible in Table 5.

Table 7: Seo and Shin (2016) FD-GMM dynamic estimator, instrumented

Thresholds (I)	g_{it-1} (II)	g_{it-1} (III)
Trimming	0.02	0.2
k	56227$***	48771$***
	(767)	(4311)
${u_{it-1}}^{cycleL}$	-0.019	0.004
	(0.025)	(0.049)
${s_{it-1}}^{cycleL}$	-0.287	-0.078
	(0.276)	(0.140)
${g_{it-1}}^{cycleL}$	0.785	0.367
	(0.175)	(0.288)
${u_{it-1}}^{cycleH}$	0.126	-0.607*
	(1.660)	(0.358)
${s_{it-1}}^{cycleH}$	3.479	-0.553**
	(5.056)	(0.310)
${g_{it-1}}^{cycleH}$	-17.924	1.793*
	(13.476)	(0.959)
Instrument	No	Yes
T	22	22
N	15	15

*** Null rejection at 1% level; ** Null rejection at 5% level; * Null rejection at 10% level. Country fixed-effects included. Column (I): relevant statistics; Column (II),(III): estimated parameters for the indicated threshold variable value. The apex values *cycleL* and *cycleH* indicate that the cyclical regressor component belongs to either the regime below (L) or above (H) the threshold. Standard Errors in parentheses. Objective variable: unemployment rate, u_{it}.

[18]We tested all possible trimming windows from 10% to 1%: results stayed absolutely consistent with the values reported in Column (II) of table 7 and are available in Table 8. The upward jump on the average threshold value can be mainly explained by the composition and higher relative GDP per capita of the subset of countries available for instrumentation.

Table 8: Threshold selection, instrumented model

Trimming (I)	k (II)	$CI-$ (III)	$CI+$ (IV)	P-values (V)
0.01	52968$	47169$	58766$	0.001
0.02	48771$*	40321$	57221$	0.001
0.03	48775$	40247$	57304$	0.001
0.04	48718$	40278$	57160$	0.001
0.05	48781$	40330$	57233$	0.001
0.10	48740$	40295$	57185$	0.001

* Selected threshold. Column (I): trimming value as a proportion of the whole series; Column (II): threshold value; Column (III) upper confidence interval; Column (IV) lower confidence interval.

3.6. Gonzalez et al. (2017) Fixed effect logistic smooth transition model

In the previous section, the use of FD-GMM estimator allowed us to correct for both the magnitude of the GDP per capita threshold and the relative magnitude (and thus importance) of the cyclical parameters of self-employment and unemployment, drawing a comparison showing how labor market inertia and entrepreneurial spirit affect unemployment across the selected threshold.

There is however a missing piece in the puzzle: the likelihood of such transition to take place in a given neighborhood of the exact threshold value, conditional on the speed of change of the threshold variable. In the section dedicated to nonlinearity identification and static Fixed effects estimates, we detected two possible competing k values for the GDP per capita threshold. Our hypothesis is that such sharp difference might be capture by a model capable of smoothing out the regression across a dynamic threshold.

To keep the analysis consistent, we will focus again on both our candidate thresholds. In order to account for a potential, non-sudden transition, we follow Gonzalez et. al and employ a Lagrangeian Multiplier test that compares the 'null' of linearity with a smooth, nonlinear transition multiple regime model. The authors suggest following Luukkonen, Saikkonen and Teräsvirta (1988), which start from a first order Taylor expansion of the transition function $f(d_{it-1}; \gamma, k)$ around its slope γ, setting its starting value to 0. After a reparametrization of the model, we would thus be left with an auxiliary regression function testing the null $H_0 : \gamma = 0$ and amounting to:

$$H_0^i : \beta_i = \beta_{i-1} = \beta_{i-2} = \ldots = \beta_1 = 0 \tag{18}$$

for any possible $i = 1, ..., m$ where m is the maximum number o threshold, and the β_i coefficients indicate the different parameter values of the regressors, excluding the nonlinear ones. Following the results from the Hansen (1999) linearity test, we set "i" to 1. Results for the LM tests are visible in Table 9. The null of linearity is strongly rejected in favor of the statistical significance of the polynomial Taylor expansion, thus allowing to reject strongly the original null hypothesis $H_0 : \gamma = 0$. That said, we are now left with the estimate of the optimal logistic transition function, including its regime parameters, the slope of the logistic and the threshold value for both GDP per capita and the unemployment rate. We remind that, based on the validity of previous estimates, our thesis is that the GDP per capita, as the "best performing" transition threshold, and given the results of the Hansen (1999) procedure, might present, albeit weakly, some form of smooth transition.

Table 9: LM tests, Terasvirta et al.(1988)

	Variable	LM_X	$P-value$	LM_F	$P-value$
(I)	(II)	(III)	(IV)	(IV(	(VI)
$0vs1$	g_{it-1}	15.450***	0.001	4.920***	0.002
$0vs1$	u_{it-1}	13.370***	0.004	4.255***	0.005

** Null rejection at 1% level. Lagrangean multiplier tests

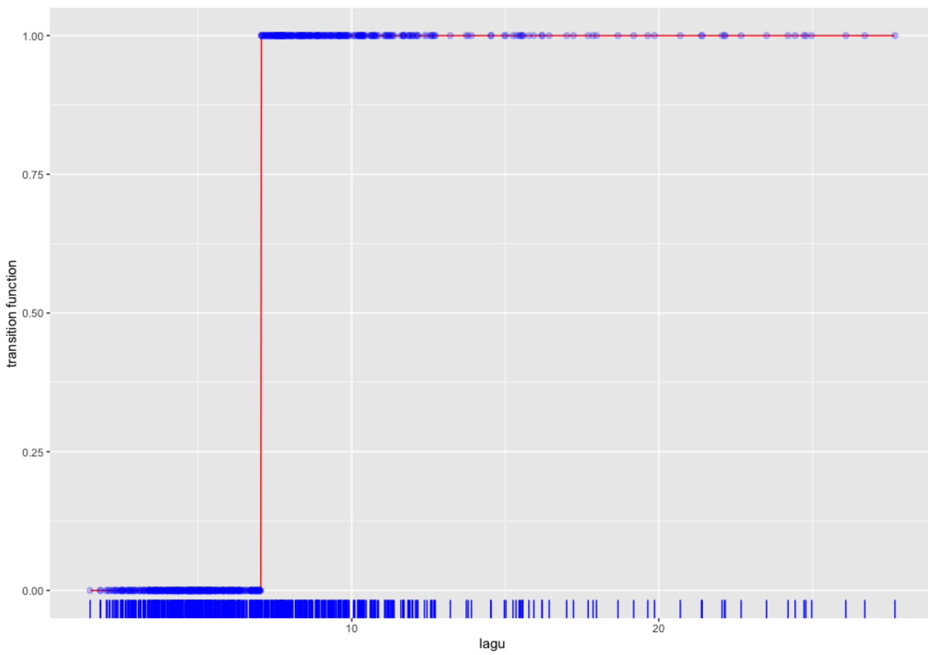
Table 10: Gonzalez et al. (2017) Panel Smooth transition regression

$Thresholds$	g_{it-1}	u_{it-1}
(I)	(II)	(III)
k	43210$***	7.052%***
	(1008$)	(0.001%)
γ	0.003***	1660***
	(0.001)	(69)
${u_{it-1}}^{cycleL}$	0.218***	0.067
	(0.057)	(0.090)
${s_{it-1}}^{cycleL}$	-0.055	-0.179***
	(0.103)	(0.088)
${g_{it-1}}^{cycleL}$	-23.670***	-15.550***
	(5.844)	(5.393)
${u_{it-1}}^{cycleM}$	0.071	0.307***
	(0.076)	(0.1408)
${s_{it-1}}^{cycleM}$	-0.310	0.047
	(0.203)	(0.179)
${g_{it-1}}^{cycleM}$	14.060***	1.221
	(5.514)	(10.400)
${u_{it-1}}^{cycleH}$	0.289***	0.374***
	(0.068)	(0.078)
${s_{it-1}}^{cycleH}$	-0.365***	-0.132
	(0.151)	(0.156)
${g_{it-1}}^{cycleH}$	-9.616***	-14.330***
	(2.511)	(6.836)
T	28	28
N	23	23

*** Null rejection at 1% level; ** Null rejection at 5% level; * Null rejection at 10% level. Country fixed-effects included. Column (I): relevant statistics; Column (II),(III): estimated parameters for the indicated threshold variable value. The apex values *cycleL*, *cycleM* and *cycleH* indicate that the cyclical regressor component belongs to either the regime below the threshold (L), the intermediate nonlinear regime (M) or to the regime above the threshold (H). Standard Errors in parentheses. Objective variable: cyclical unemployment rate, ${u_{it}}^{c}$.

Estimates for the optimized models are available in Table 10. We will focus once more our attention on GDP per capita as a threshold: although we kept testing for unemployment as a threshold, and albeit the threshold value could be estimated precisely enough at 7.052%, the sharpness of the transition ($\gamma = 1660$) suggested a perfectly discrete jump between regimes, so that a smooth, or even remotely smooth transition function was not achievable, even after quite a few rounds of optimization. An illustration of the tentative transition function for $d_{it-1} = u_{it-1}$ can be seen in Figure 4.

Figure 4: Logistic transition function for $u_{i,t-1}$



Let´s now focus on $d_{it-1} = g_{it-1}$. After a first round of optimization, where the starting value for the logistic was set close to 0 for both the slope and the threshold value, we are left with the following estimates, $\gamma = 0.0002$ and $k = 43210\$$. We use such final values again to rerun the algorithm, focusing on increasing the value of γ in order to find the best fit for the model, gradually leaving the standard threshold fixed effect ($\gamma = 0$). After some rounds, We identify the most significant slope parameter in the neighborhood $\gamma = (0.1, 0.3)$.[19] Table 10 thus reports results of the model for $\gamma = 0.15$ and $k = 43210\$$.

As we can see, in column (II), the transitional effect is still present, and the final value of the threshold finds itself in-between the estimates for the Fixed effect Hansen (1999) type estimator and the FD-GMM estimates of Seo and Shin (2016). The optimized threshold value also appear to be the closest to the panel average. A representation of the smoothing function for the threshold value with respect to the observed GDP per capita values is visible in Figure 5, while its corresponding, nonlinear parameter variation is available in Figures 6, 7 and 8.[20]

[19]Estimated coefficient for the values across the 0.1 - 0.3 interval all brought similar results in terms of magnitude and sign.

[20]In Figure 9 we also offer an illustration of the response (value) of the logistic smooth function given different threshold values and the corresponding estimate of the self-employment coefficient.

Figure 5: Logistic transition function for $g_{i,t-1}$

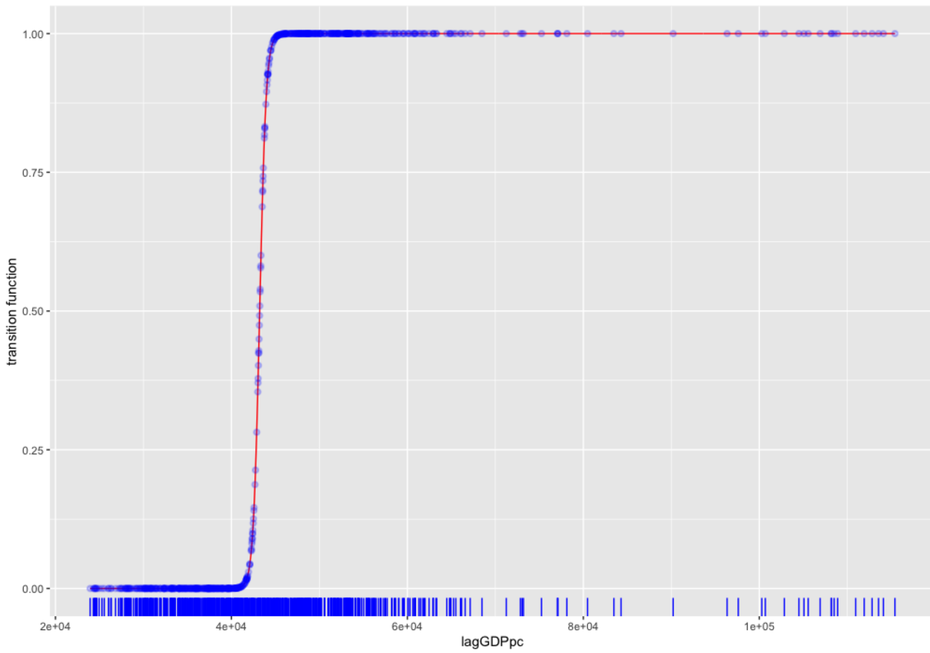


Figure 6: Coefficient transition for ${g_{i,t-1}}^c$

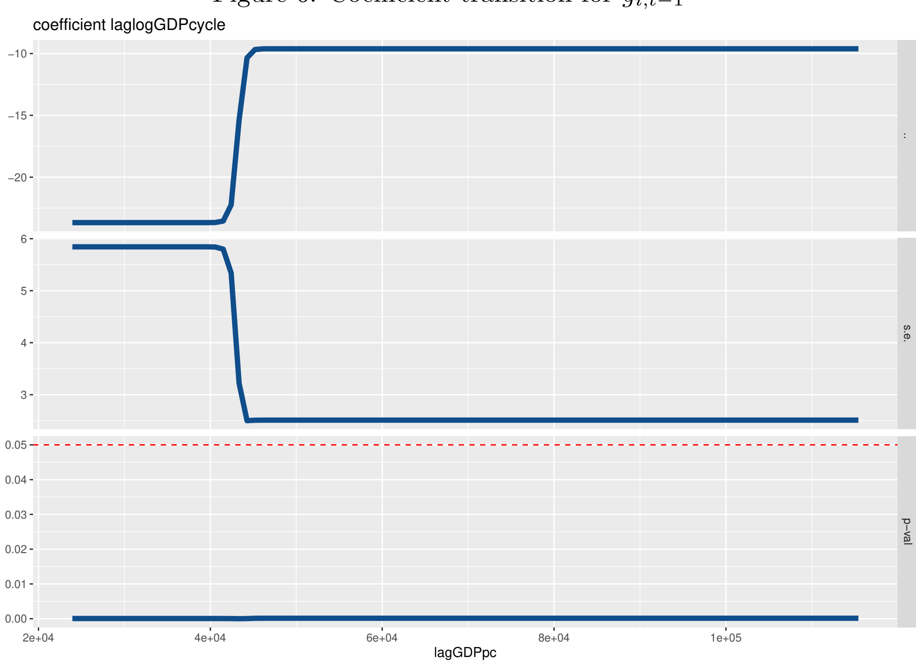


Figure 7: Coefficient transition for ${u_{i,t-1}}^c$

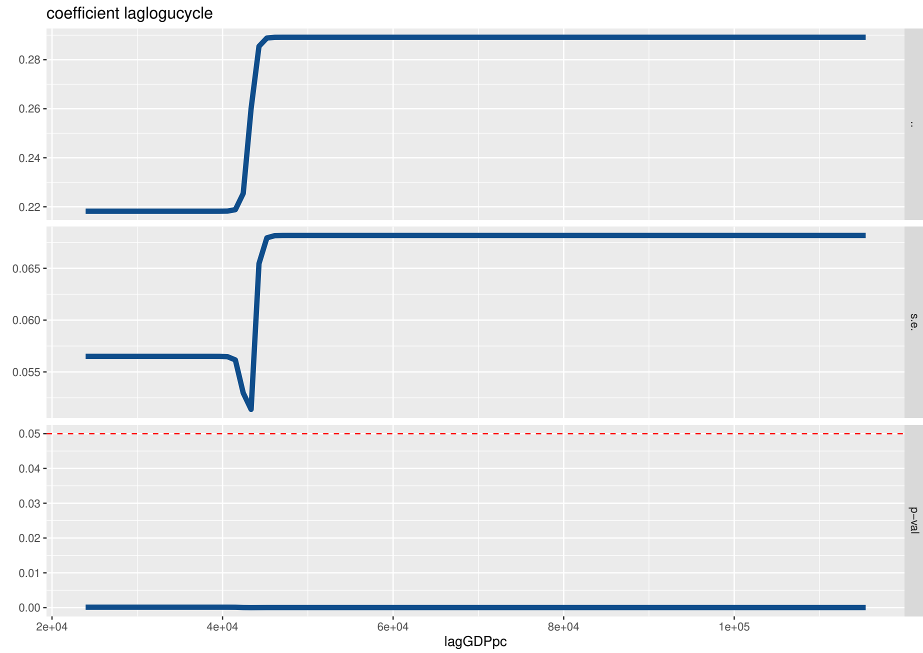
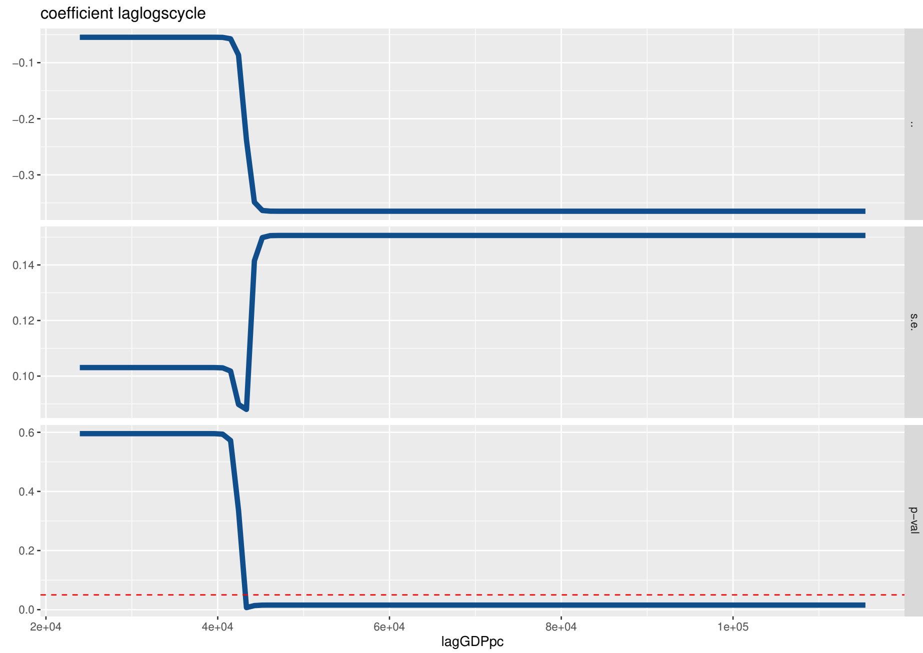


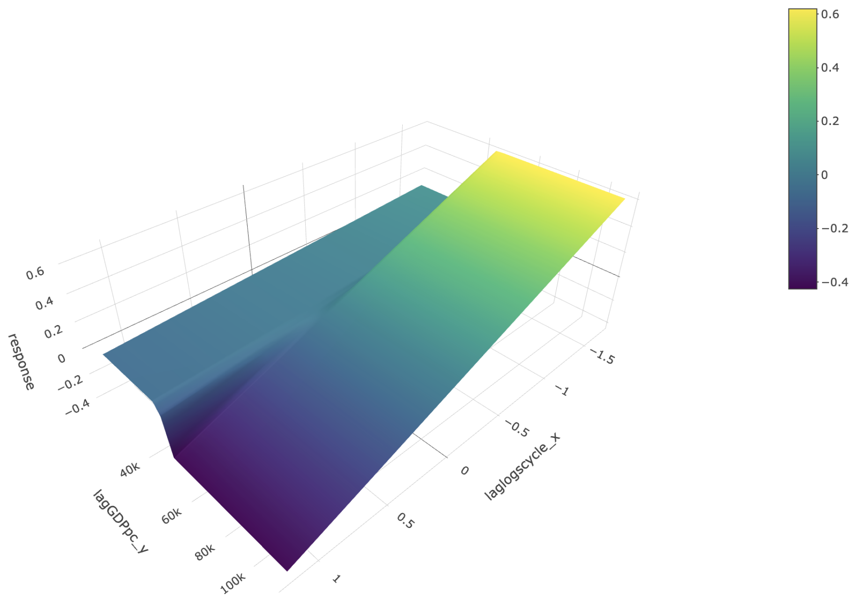
Figure 8: Coefficient transition for ${s_{i,t-1}}^c$



The nature of the results consistently follow the estimates of the non dynamic and dynamic fixed effects threshold estimators. Once we enter the upper regime, self-employment negatively enters the relationship, although with some delay, identifiable in the immediate neighborhood of the estimated GDP per capita threshold value. As the nonlinear estimates do not take the upward bias of the autoregressive unemployment coefficient into account, the very relevant change in the relative magnitude between the coefficient for ${u_{i,t-1}}^c$ and ${s_{i,t-1}}^c$ is once again lost. Without any doubt, we observe how the regime switch did not contribute much to change the inertia of unemployment (the ${u_{i,t-1}}^c$ coefficient just raised by 0.071 in the neighborhood of the threshold) but added very significatively to the relationship between cyclical unemployment and cyclical self-employment, weighting more than a six-fold than the lower regime in the higher one, after the transition has been completed (coefficient of ${u_{i,t-1}}^c = -0.365$).

In this case, the results regarding the relationship, using the aggregate self-employment rate, merely confirm the previous findings while also providing evidence of the relationship in the medium regime, not just in the high regime.

Figure 9: Surface response graph for ${s_{i,t-1}}^c$



3.7. An alternative metric for entrepreneurship using the Kremer, Bick and Nautz estimator (2012)

To verify whether our results may be affected by the aggregate entrepreneurship metric adopted, in this section we present the findings obtained using the TEA measure, specifically nascent entrepreneurship, sourced from the GEM database. This option presents an additional challenge if we want to continue working with panel threshold estimators due to the nature of the short panel. This challenge is addressed through the use of a more flexible estimator: the Kremer, Bick, and Nautz (2012) estimator.

We now switch our definition of entrepreneurship, focusing on the so-called nascent entrepreneur (n_{it}) the way it has been defined by the Global Entrepreneurship Monitor (GEM). In particular, we shall focus on the The nascent entrepreneurship rate as a variable measuring...

> ...when one or more persons start to commit time and resources to founding a new firm. If they do so on their own, and if the new venture can be considered as an independent start-up, they are called nascent entrepreneurs.

The GEM puts at our disposal a measure of nascent entrepreneurial intensity normalized as a percentage of the total workforce.[21] The data out our disposal spans the 1992-2020 period for a group of around 13 countries. After controlling for the availability of the exogenous instrument we have been using, such number naturally decreases.

Unfortunately, given the limited size of the database, computations of the models conditional on the availability of the original instrumental variable in each time period for each cross sectional unit where not possible: thus, as far as the Seo and Shin estimator goes, we can only provide the estimates for the standard GMM-FD estimator and the more refined estimates augmented with additional lags of the endogenous regressors, n_{it} and u_{it}.

Table 12 gives us an overview of the possible threshold value option for our three specifications. As we would expect, although performing surprisingly well until a cut off equivalent to a 0.04 trimming of

[21] Together with a second ratio which relates nascent entrepreneurs to the total of self entrepreneurs, thus including nascent intrapreneurs as well.

the time dimension of each cross section, the threshold estimates for the standard model in Column I quickly collapses to a different and non significant threshold value at 0.05. As we already observed with the business creation measure, taking into account endogeneity through additional instrumentation of the regressors does help the procedural identification of the threshold: columns III and IV show a much more stable and significant value across all possible trimming levels[22].

Final estimates for the models cited in this section are finally available in Table 11. On a positive note, we observe how business ownership and nascent entrepreneurship do differ in terms of the threshold governing the transition across states: the low, 15000$ confirmed threshold in the previous section has now left space to a much higher value, in the neighborhood of 42000$.

On the negative side, the use of the GEM database, conditional on the necessity to look for a way to control for endogeneity, comes with a strong limitation: as we observe Table 11, we will find hardly any parameter to be statistically relevant, and instrumental lags to make hardly any difference in the GMM estimation across the three models presented: this is due to the disproportion between the N and the T dimension, as the latter now amounts to more than a threefold of the former, thus making the estimates of the transition highly imprecise[23].

Table 11: Seo and Shin (2016), GEM Estimates

(I)	(II)	(III)	(IV)
k	42369$***	42369$***	42369$***
	(11175)	(11173)	(11173)
${u_{it-1}}^{cycleL}$	-0.053	-0.053	-0.053
	(3.336)	(3.336)	(3.336)
${n_{it-1}}^{cycleL}$	-1.404	-1.404	-1.404
	(4.184)	(4.184)	(4.184)
${g_{it-1}}^{cycleL}$	-2.977	-2.977	-2.977
	(8.190)	(8.190)	(8.190)
${u_{it-1}}^{cycleH}$	-1.493	-1.493	-1.493
	(2.980)	(2.980)	(2.980)
${n_{it-1}}^{cycleH}$	2.668	2.668	2.668
	(19.478)	(19.478)	(19.478)
${g_{it-1}}^{cycleH}$	14.818	14.818	14.818
	(9.661)	(9.661)	(9.661)
Instrument	no (dep. only)	u_{it-1}	u_{it-1} and n_{it-1}
T	28	28	28
N	9	9	9

*** Null rejection at 1% level; ** Null rejection at 5% level; * Null rejection at 10% level. Country fixed-effects included. Column (I): relevant statistics; Column (2),(3), (4): estimated parameters for the indicated threshold variable value. The apex values *cycleL* and *cycleH* indicate that the cyclical regressor component belongs to either the regime below (L) or above (H) the threshold. Standard Errors in parentheses. Trimming value of the T dimension: 0.04. Objective variable: unemployment rate, u_{it}.

[22] As for any other exercise of the sort contained in this draft, trimmings were considered up to the maximum, literature-relevant limit of 15% (0.15). No relevant changes in the above considered patterns were detected, apart from the impossibility of the regression window to compute correctly the threshold value at 0.15 given the far too small number of observations left after the trimming procedure

[23] The Seo and Shin (2016) estimator still relies on asymptotics that require the N dimension to converge faster to infinity than the T dimension: although one could get away with a higher number of time periods in an empirical application, a strong disproportion between the two dimensions is very likely to cause precision issues, as in our case. The issue, we need to point out, did not change when considering the rate of nascent entrepreneurs calculated with respect to the total of self-employment (That is, considering intrapreneurship as well and standardizing nascent entrepreneurs by the self-employment total).

Table 12: GEM, Threshold selection

Trimming (I)	dep. only (II)	u_{it-1} (III)	u_{it-1} and n_{it-1} (IV)
0.01	42383$***	42383$***	42383$***
0.02	42394$***	42393$***	42393$***
0.03	42385$***	42384$***	42384$***
0.04	42369$***	42369$***	42369$***
0.05	39516$	42374$***	42375$***

Threshold value for the range $0.01 - 0.05$. Column (I): trimming value as a proportion of the whole series; Column (II): instrument: lag of dependent variable; Column (III) instruments: lag of unemployment; Column (IV): instrument: lag of unsemployment and nascent entrepreneurship. ***:significant at 1%; **: significant at 5%; *: significant at 10%.

The unsatisfactory results from the Seo and Shin (2016) estimates shown above, which lead to very low statistical precision, can be explained in terms of: the limited amount of total observations at our disposal, which limits both the grid-search (number of repetitions in the algorithm estimating the true value of the threshold) and puts strong limitations on the trimming value (which is limited between 0.01 and 0.05, being fairly distant from the 0.10 - 0.15 upper limit established by practitioners); the asymptotic limit given by the theory related to the Hansen (1999) estimator and all the most recent ones derived from it, favoring the N dimension rather than the T dimension; the possible existence of some patterns of heterogeneity in the self-employment measures (say, pull and push factors differentiating the measure) that might be affecting the estimates.

In relationship to the three limitations defined above, this section presents, in Table 13, the estimates from the Kremer, Bick and Nautz (2012) dynamic panel threshold estimator.

Table 13 reports two different normalizations for nascent entrepreneurship, available in the GEM database: $\frac{n}{wf}$% as the ratio between nascent entrepreneurs and total work-force; $\frac{n}{s}$% as the ratio between nascent entrepreneurs and self-employed. In both cases, as the main body of our research concluded that the lagged value of GDP would perform better in nonlinearity testing and as such be more apt at capturing nonlinearities in the models, we report g_{it-1} as the only feasible regime switching indicator.

Focusing on our results, the implementation of the Kremer, Bick adn Nautz (2012) estimator, keeping in mind advantages and limitations already highlighted with respect to those seen in the main analysis, shows results which fall, to some extent, in line with those seen in the main text. In the lower regime, when g_{it-1} is taken as a Threshold variable[24], n_{it-1} appears disconnected from cycle: in the upper regime, the relative statistical significance of the variable picks up, and we are left with a pattern of strong dependence on national health.

We would also need to remember how delicate the procedure might be when a threshold value is not uniquely identified even if the modelization imposes a two-regimes structure. In Figures 10 and 11 we report the estimate for the γ value for the threshold models in Columns (I) amd (II) of Table (13): for GDP, a unique value confirming the threshold behavior of the model is identified. On the contrary, when unemployment is taken as a threshold, two potential local values, both capable of rejecting the mull of linearity and very close one to the others, are detected by the recursive estimates of the model, impacting the precision of the estimated coefficients.

[24]We again tested for u_{it-1} as a potential threshold with the linearity test, and found out once again that the null of linearity could not be rejected

Table 13: Kremer, Bick and Nautz (2012), GEM Estimates

(I)	(II)	(III)
n definition	$\frac{n}{s}\%$	$\frac{n}{wf}\%$
k	37441$***	37441$***
$C.I.low$	30642	30642
$C.I.high$	48904	48904
u_{it-1}	1.108***	1.077***
	(0.081)	(0.095)
g_{it-1}	0.001**	-0.001
	(0.000)	(0.000)
${n_{it-1}}^{cycleL}$	-0.211	-0.733
	(0.248)	(0.461)
${n_{it-1}}^{cycleH}$	-0.070**	-0.533***
	(0.029)	(0.124)
$Threshold$	g_{it-1}	g_{it-1}
$Instruments$	u_{it}, $protestants$	u_{it}, $protestants$
T	21	21
N	7	7

*** Null rejection at 1% level; ** Null rejection at 5% level; * Null rejection at 10% level. Column (I): relevant definitions; Columns (II),(III): estimated parameters for the indicated threshold variable value. The apex values $cycleL$ and $cycleH$ indicate that the cyclical nascent entrepreneur parameter belongs to either the regime below (L) or above (H) the threshold. Standard Errors in parentheses. Trimming value of the T dimension: 0.10. Objective variable: unemployment rate, u_{it}. Up to 4 lags of the instrument $protestants$ were evaluated to enter the first step regression of the 2SLS procedure after differencing. Up to 4 lags of the instruments u_{it} were evaluated to enter the second step (level) regression of the 2SLS procedure after differencing. Country-specific sfixed effects included.

Figure 10: Value of γ, GDP

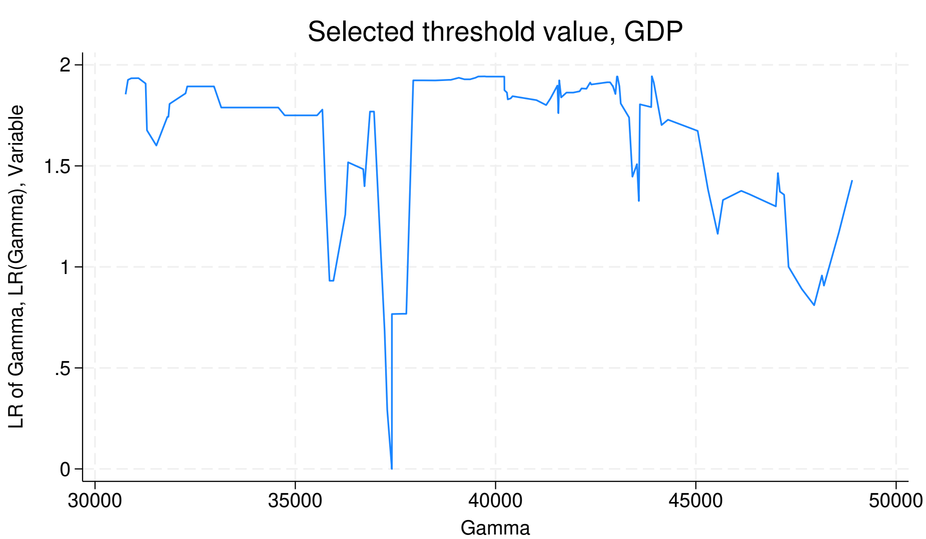


Figure 11: Value of γ, unemployment

The additional modelization above adds to the robustness of the results, but it is still, considering the 147 total observations at our disposal and the three-to-one ratio of the dimensions of the panel, a bit far away from being totally conclusive. As aggregation techniques to obtain longer time series and, perhaps more importantly in our framework, additional country enter the GEM granting a balanced panel dataset, additional evidence will be needed in the future.

In sum, there is statistically significant evidence of a contribution from new entrepreneurs to unemployment reduction, but only in economies classified under the high economic performance regime (defined as having a per capita income level above the threshold). This finding is robust regardless of the metric used (self-employment/nascent entrepreneur), the database employed (ILO/GEM), the threshold panel model utilized, and the treatment of reverse causality.

3.8. Exploring heterogeneity: Start-up motive

The composition of the self-employment sector can significantly influence the dynamics of the entrepreneurship-unemployment nexus, highlighting the need for a nuanced understanding of how different types of self-employment contribute to overall economic outcomes, particularly job creation and labor market dynamics. Although most empirical studies seem to have overlooked this issue, some works have considered the heterogeneity of self-employment from various perspectives. Apergis (2015) focuses on identifying potential differences in this relationship by utilizing disaggregated self-employment rates that distinguish between incorporated and unincorporated businesses. Meanwhile, Haussen and Schlegel (2022) explore gender differences in this relationship and examine the participation of solo self-employed workers in new entries to self-employment. Finally, Apergis and Payne (2016), revisit empirically the relationship using the Kaufmann Entrepreneurship Index for U.S. states, a metric that aims to provide a better approximation of entrepreneurship measurement than the classical operationalization through self-employment.

In parallel to this last group of studies, we aspire to contribute to a better understanding of this relationship by exploring a source of heterogeneity that can potentially affect the employment intensity of self-employment: the motivation behind becoming an entrepreneur. The startup motive is closely related to the business economic performance of a venture, one of whose dimensions is its capacity to create employment (Caliendo et al., 2014). Therefore, controlling for the startup motive equates to recognizing that not all new self-employment is entrepreneurial. In this way, we investigate the idea that not all self-employment is inherently entrepreneurial (true entrepreneurs), aiming to isolate the contribution of this type of self-employment to job creation in contrast to “non-entrepreneurial” forms.

At this point, we hypothesize that much of the previous literature has likely overlooked the fact that not all self-employment contributes in the same way or under the same circumstances to the reduction of unemployment. In particular, we argue that depending on the type of self-employment (by start-up motive) driving the change in self-employment rates (opportunity-based versus necessity-based entrepreneurs), the impact on the unemployment rate can vary, depending on their level of economic performance.

To this purpose, we exploit what the GEM database currently has to offer in order to tackle additional sources of heterogeneity coming from pull and push entrepreneurial definitions available in the database. We employ two entrepreneurial definitions based on the concept of necessity and opportunity: nascent opportunity workers over the total of self-employed ($\frac{n_o}{s}\%$), and nascent necessity workers over the total of self-employed (($\frac{n_n}{s}\%$). Results for the four measures are visible in Table 14, evaluated with respect to the threshold variable g_{it-1} and estimated once again with the Kremer, Bick and Nautz (2012) estimator.[25]

The heterogeneity the two measures embody grant us some additional insights over the results shown in the main body of our work: the first effect, related to entrepreneurial creation, is visible in Column (III) and can be directly compared the final results from the Shin and Seo (2016) estimates. For GDP values above a given threshold, the negative relationship between unemployment and necessity self employment kicks in.

On the other side, a similar effect is given by nascent pull entrepreneurship: although the effect for nascent opportunity entrepreneurs appears to be weaker, it is statistically precise (Column (II)), with the obtained results implying a negative relationship between Nascent opportunity entrepreneurship and unemployment below a certain GDP threshold. This would suggest that nascent opportunity entrepreneurship act as a job-creating force in those countries characterized by a situation of lower economic performance.

Whatever the case, table 14 denotes a strong, negative net effect (-0.701 in column (III) compared to -0.216 in column (II)) on the necessity entrepreneurship side when countries face time periods where GDP is standing above the threshold value, in the upper regime.

In sum, the finding is that there is a systematic relationship between self-employment once we control for the startup motive, but regime-dependent. Specifically, we categorize self-employed individuals based on their start-up motive (opportunity vs. necessity entrepreneurs). We then investigate whether changes in new entries of opportunity or necessity entrepreneurs drive unemployment reductions, and whether these effects are different depending on the level of economic performance. The results suggest that reductions in unemployment are driven by positive changes in the proportion of opportunity entrepreneurs when a country is operating below a certain income threshold (i.e., in low performing-countries), whereas changes in the participation of push-factor or necessity-based entrepreneurship lead to significant reductions when a country is above that threshold (in higher-performing countries).

This has several implications and potential explanations. First, in economies within the first regime, promoting opportunity-driven entrepreneurs proves to be effective in reducing unemployment. This result points to the importance of having an appropriate ecosystem and incentive structure that supports the emergence of these ”true entrepreneurs” for job creation, in these contexts. These strategies not only foster economic growth but also contribute to a decrease in the unemployment rate. One possible interpretation is that if an economy is positioned below the threshold, there is room to qualify entrepreneurship, thus transforming society into a more entrepreneurial one and enhancing the employment intensity of self-employment. This entails fostering entrepreneurship and ensuring that new entries consist of scalable businesses with a higher likelihood of becoming job creators.

In contrast, in economies within the second regime, opportunity-based entrepreneurs do not seem

[25] Notice that we have abstracted one again from presenting the results with u_{it-1} as a threshold: in none of the four specifications of the table such variable was found to be a valid threshold or resulted in statistically significant estimates.

to significantly impact unemployment reduction. This may be due to increased competition and the generally smaller average size of new firms. However, variations in the share of necessity-based entrepreneurs do appear to reduce unemployment. One can hypothesize that in these economies—some of which may be nearing full employment—the opportunity cost of self-employment for a significant number of skilled workers or those with in-demand skills may be elevated, as many of the most lucrative business opportunities may have already been captured. In such economies, self-employment may serve as a refuge for workers with lower employability, who will account for a substantial portion of the growth in newly created jobs. This process may be further accelerated by emerging trends in the labor market, such as the rise and increasing penetration of on-demand employment characteristic of the gig sector. While this effect may not be entirely desirable, it can help workers with lower employability and those marginally attached to the labor market gain employment. On the downside, a large portion of these jobs are associated with dependent forms of self-employment, often precarious, and with the substitution of traditional salaried jobs for flexible work arrangements —such as those in the gig economy.

Table 14: Kremer, Bick and Nautz (2012), Nascent Opportunity and Nascent Necessity Entrepreneurs, GEM Estimates

(I)	(II)	(III)
n definition	$\frac{n_o}{s}\%$	$\frac{n_n}{s}\%$
k	30904$***	48903%***
$C.I.low$	30642	30642
$C.I.high$	48903	48903
u_{it-1}	0.982***	0.867***
	(0.049)	(0.092)
g_{it-1}	0.001	-0.001
	(0.000)	(0.000)
${n_{it-1}}^{cycleL}$	-0.216**	-0.175
	(0.110)	(0.115)
${n_{it-1}}^{cycleH}$	-0.005	-0.701***
	(0.025)	(0.124)
$Threshold$	g_{it-1}	g_{it-1}
$Instruments$	u_{it}, $protestants$	u_{it}, $protestants$
T	21	21
N	7	7

*** Null rejection at 1% level; ** Null rejection at 5% level; * Null rejection at 10% level. Column (I): relevant definitions; Columns (II),(III): estimated parameters for the entrepreneural definition. The apex values *cycleL* and *cycleH* indicate that the cyclical nascent entrepreneur parameter belongs to either the regime below (L) or above (H) the threshold. Standard Errors in parentheses. Trimming value of the T dimension: 0.10. Objective variable: unemployment rate, u_{it}. Up to 4 lags of the instrument *protestants* were evaluated to enter the first step regression of the 2SLS procedure after differencing. Up to 4 lags of the instruments u_{it} were evaluated to enter the second step (level) regression of the 2SLS procedure after differencing. Country specific fixed effects included.

4. Conclusions

This study examined the intricate relationship between entrepreneurship and unemployment, focusing on key issues such as non-linearity, reverse causality, and the heterogeneity of self-employment types. By applying panel threshold estimators, the research offered a nuanced understanding of how different forms of entrepreneurship—specifically opportunity and necessity entrepreneurship—affect unemployment rates across various economic contexts. The findings reveal shortcomings in existing literature that often assumes a linear and systematic relationship between aggregate self-employment and unemployment. Our initial results indicate that such a relationship is only present in economies characterized by high economic performance.

The contributions of this work are critical to enhancing our understanding of the entrepreneurship-unemployment nexus. Previous studies have frequently overlooked the complexities involved in this relationship, primarily focusing on aggregate metrics without considering the underlying dynamics that differentiate types of entrepreneurship. Our findings demonstrate that the dynamics of self-employment are not uniform. The analysis of startup motives highlights that the economic impact of entrepreneurship is contingent on a country's economic performance. Specifically, the results showcase the essential role of startup motives: opportunity entrepreneurs drive unemployment reduction in low-performing economies, while necessity entrepreneurs significantly impact high-performing contexts. This interplay between startup motives and economic performance underscores the necessity of considering the motivations behind self-employment to gain clearer insights into how different types contribute to overall economic outcomes, especially job creation.

Although primarily focused on measurement, this paper carries practical implications for policymakers. A nuanced understanding of the differentiated impact of self-employment types can inform strategies that promote entrepreneurship as a means to address unemployment challenges. This is particularly relevant in light of recent trends in self-employment observed in leading economies, including the notable growth of the gig economy. Additionally, the study opens avenues for future research, especially regarding the quality of jobs created through entrepreneurship and the evolving dynamics of labor markets influenced by economic conditions. By further exploring these aspects, scholars can refine their understanding of the complex relationship between entrepreneurship and unemployment in a rapidly changing economic landscape.

In summary, this research offers valuable contributions to the literature, calling for a re-evaluation of assumptions surrounding self-employment's effects on unemployment rates. The findings are especially pertinent considering the recent evolution of self-employment in leading economies. As emphasized by Boeri et al. (2020) and Henley (2022), there has been a significant rise in self-employment, particularly due to economic shifts associated with digital platforms and gig work. This trend reflects not only the changing nature of work but also the growing importance of entrepreneurial activity in job creation. Recent data indicate that a substantial portion of job growth in leading economies is linked to self-employment, encompassing various forms of gig work that offer flexibility but also present challenges in terms of job security and benefits. The dynamics of necessity and opportunity entrepreneurship are closely associated with these trends, highlighting the need for policymakers to recognize and differentiate between these two paths. By doing so, they can better address the unique challenges arising from the gig economy and ensure that self-employment contributes positively to economic development.

This study encourages several promising avenues for future research, particularly related to the nuanced relationship between entrepreneurship and unemployment. One critical area involves measuring not only the contribution of changes in self-employment rates to unemployment reduction but also the quality of the employment generated. This focus is increasingly relevant given that much of the employment growth post-Great Recession has been driven by atypical forms of self-employment, which often replace traditional employment relationships and can be associated with precariousness and underemployment (Borowczyk-Martins & Lalé 2020; Bell & Blanchflower 2021; Congregado et al. 2024b).

Moving beyond aggregate analyses that fail to distinguish between different types of entrepreneurs is essential. This study underscores the importance of examining the distinctions between necessity and opportunity entrepreneurs, as well as exploring new dimensions such as the voluntary nature of self-employment and the phenomenon of underemployment. Utilizing alternative measures of unemployment and self-employment, including involuntary part-time employment and marginally attached workers, will enhance our understanding of these dynamics. Such an approach can explain the varied findings while reflecting new trends in labor markets, particularly the rapid rise of contracting and atypical self-employment arrangements. Leveraging Labor Force Surveys for identifying these dimensions can significantly improve our understanding of the dynamic relationship between entrepreneurship and unemployment.

5. References

Aarstad, J., Kvitastein, O.A., Jakobsem, S.E., 2016. Related and unrelated variety as regional drivers of enterprise productivity and innovation: A multilevel study. Res. Policy 45(4), 844-856. https://doi.org/10.1016/j.respol.2016.01.013

Acs, Z.J., 2006. How Is Entrepreneurship Good for Economic Growth? Innovations: Technology, Governance, Globalization 1(1), 97-107. http://dx.doi.org/10.1162/itgg.2006.1.1.97.

Acs, Z.J., Amoros, J., 2008. Entrepreneurship and competitiveness dynamics in Latin America. Small Bus. Econ. 31, 305-322. http://dx.doi.org/10.1007/s11187-008-9133-y

Acs, Z.J., Armington, C., 2004. Employment Growth and Entrepreneurial Activity in Cities. Reg. Stud. 38, 911-927. http://dx.doi.org/10.1080/0034340042000280938

Acs, Z.J., Armington, C., 2002. The Determinants of Regional Variation in New Firm Formation. Reg. Stud., 36, 33-45. https://doi.org/10.1080/00343400120099843

Acs, Z.J., Varga, A., 2005. Entrepreneurship, agglomeration and technological change. Small Bus. Econ. 24, 323-334. http://dx.doi.org/10.1007/s11187-005-1998-4

Amoros, E., Cristi, O., Minniti, M., 2009. Driving Forces behind entrepreneurship: differences on entrepreneurship rate levels and its volatility across countries, Frontiers in Entrepreneurship Research 29, Issue 16, article 2. http://digitalknowledge.babson.edu/fer/vol29/iss16/2

Arin, K.P., Huang, V.Z., Minniti, M., Nandialath, A.M., Reich, O.F., 2015. Revisiting the determinants of entrepreneurship: A Bayesian approach. J. Manage. 41(2), 607-631. http://dx.doi.org/10.1177/0149206314558488

Armington, C., Acs, Z.J., 2004. Job creation and persistence in services and manufacturing. J. Evol. Econ. 14, 309-325. http://dx.doi.org/10.1007/s00191-004-0194-z

Audretsch, D.B., Dohse, D., Niebuhr, A., 2015. Regional unemployment structure and new firm formation. Pap. Reg. Sci. 94, S115-S138. http://dx.doi.org/10.1111/pirs.12169

Audretsch, D., Thurik, R., 2004. A model of the entrepreneurial economy. International Journal of Entrepreneurship Education 2, 143-166. http://hdl.handle.net/2262/104291

Apergis, N., Payne, J.E., 2016. An empirical note on entrepreneurship and unemployment: Further evidence from US States. Journal of Entrepreneurship and Public Policy, 5(1), 73-81. https://doi.org/10.1108/JEPP-10-2015-0029

Baptista, R., Escaria, V., Madruga, P., 2008. Entrepreneurship, regional development and job creation: the case of Portugal. Small Bus. Econ. 30, 49-58. http://dx.doi.org/10.1007/s11187-007-9055-0

Bartelsman, E., Scarpetta S., Schivardi, F., 2005. Comparative analysis of firm demographics and survival: evidence from micro-level sources in OECD countries. Ind. Corp. Change 14(3), 365-391. http://dx.doi.org/10.1093/icc/dth057

Baumgartner, H.J., Caliendo M., 2008. Turning Unemployment into Self-Employment: Effectiveness of Two Start-Up Programmes. Oxford B. Econ. Stat 70, 346-373. http://dx.doi.org/10.1111/j.1468-0084.2008.00505.x

Blanchflower, D.G., 2004. Self-employment: More may not be better. NBER Work. Pap. Ser. 10286. http://dx.doi.org/10.3386/w10286

Block, J.H., Sandner, P., 2009. Necessity and opportunity entrepreneurs and their duration in self-employment: evidence from German micro data. J. Ind. Compet Trade 9, 117-137. http://dx.doi.org/10.1007/s10842-007-0029-3

Boeri, T., Giupponi, G., Krueger, A. B., Machin, S., 2020. Solo self-employment and alternative work arrangements: A cross-country perspective on the changing composition of jobs. J. Econ. Perspect. 34(1), 170-195. http://dx.doi.org/10.1257/jep.34.1.170

Breitung, J., 2000. The local power of some unit root tests for panel data, in: Baltagi, B. (Ed.), Nonstationary panels, panel cointegration, and dynamic panels. Advances in Econometrics 15, 161-177. http://dx.doi.org/10.1016/S0731-9053(00)15006-6

Candelon, B. Colletaz G., Hurlin, C., 2013. Network Effects and Infrastructure Productivity in Developing Countries. Oxford B. Econ. Stat. 75, 887-913. http://dx.doi.org/10.1111/j.1468-0084.2012.00722.x

Caner, M., Hansen, B.E., 2004. Instrumental variable estimation of a threshold model. Econ. Theory 20, 813-843. http://dx.doi.org/10.1017/S0266466604205011

Card, D., Kluve, J., Weber, A., 2010. Active labour market policy evaluations: a meta-analysis. Econ. J. 120, 452–477. http://dx.doi.org/10.1111/j.1468-0297.2010.02387.x

Carmona, M., Congregado, E., Golpe. A.A., 2016. Comovement Between Self-Employment and Macroeconomic Variables: Evidence from Spain. Sage Open 2(2), 2158244012448665. https://doi.org/10.1177/2158244012448665

Carree, M.A., Thurik, A.R., 2008. The lag structure of the impact of business ownership on economic performance in OECD countries. Small Bus. Econ. 30, 101-110. http://dx.doi.org/10.1007/s11187-006-9007-0

Cima, J., Costa, J., Pinto, C., Pinto, T., 2017. The Dynamics of Self-Employment and Unemployment During the Great Recession. Int. Rev. Entre. 15(3). https://www.senatehall.com/entrepreneurship?article=579

Congregado, E. (Ed.), 2008. Measuring Entrepreneurship: Building a Statistical System, Springer, New York. https://link.springer.com/book/10.1007/978-0-387-72288-7

Congregado, E. Golpe, A.A., Carmona, M., 2010. Is it a good policy to promote self-employment for job creation? Evidence from Spain. J. Pol. Model. 32, 828-842. http://dx.doi.org/10.1016/j.jpolmod.2010.09.001

Congregado, E., Golpe, A.A., Parker, S.C., 2012. The dynamics of entrepreneurship: hysteresis, business cycles and government policy. Empir. Econ. 43, 1239-1261. http://dx.doi.org/10.1007/s00181-011-0516-6

Congregado, E., Golpe A.A., van Stel A., 2012. The 'recession-push' hypothesis reconsidered. Int. Entrep. Manag. J. 8, 325-342. http://dx.doi.org/10.1007/s11365-011-0176-1

Congregado, E., Millan, J.M., Roman, C., 2010. From own-account worker to job creator. Int. Rev. Entrep. 8, 277-30. https://www.tara.tcd.ie/bitstream/handle/2262/104533/278.pdf?sequence=1&isAllowed=y

Congregado, E., Garcia-Clemente, J., Rubino, N., Vilchez, I., 2024. Testing hysteresis for the US and UK involuntary part-time employment. Appl. Econ., 1-21. http://dx.doi.org/10.1080/00036846.2024.2302931

de Andrés, M.I., Congregado, E., Román, M.C., 2024. Precariousness in employment mediated by digital platforms. Evidence from Europe. Rev. Econ. Mund. 66, 129-149. http://dx.doi.org/10.33776/rem.vi66.8065

Cowling, M., Taylor M., Mitchell, P., 2004. Job creators. Manch. Sch. 72, 601-617. http://dx.doi.org/10.1111/j.1467-9957.2004.00411.x

Cowling, M. L., Wooden, M., 2021. Does solo self-employment serve as a 'stepping stone' to employership?. Labour Econ. 68, 101942. http://dx.doi.org/10.1016/j.labeco.2020.101942

Cueto, B., Mato, J., 2006. An analysis of self-employment subsidies with duration models. Appl. Econ. 38, 23-32. http://dx.doi.org/10.1080/00036840500367542

Fairlie, R.W., and Fossen, F.M., 2020. Defining opportunity versus necessity entrepreneurship: Two components of business creation, in: Polachek, S.W., Tatsiramos, K. (Eds.), Change at Home, in the Labor Market, and On the Job (Research in Labor Economics, Vol. 48). Emerald Publishing Limited, Leeds, pp. 253-289. https://doi.org/10.1108/S0147-912120200000048008

Faria, J.R., Cuestas, J.C., Gil-Alana, L.A., 2009. Unemployment and entrepreneurship: A cyclical relation?. Econ. Lett., 105(3), 318-320. http://dx.doi.org/10.1016/j.econlet.2009.09.004

Faria, J.R., Cuestas, J.C., Mourelle, E., 2010. Entrepreneurship and unemployment: a nonlinear bidirectional causality?. Econ. Model. 27(5), 1282-1291. http://dx.doi.org/10.1016/j.econmod.2010.01.022

Fossen, F.M., 2021. Self-employment over the business cycle in the USA: a decomposition. Small Bus. Econ. 57(4), 1837-1855. http://dx.doi.org/10.1007/s11187-020-00375-3

Fritsch, M., Mueller, P., 2004. The Effects of New Business Formation on Regional Development over Time. Reg. Stud. 38, 961-975. http://dx.doi.org/10.1007/s11187-007-9067-9

Fritsch, M., Mueller, P., 2007. The persistence of regional new business formation-activity over time – assessing the potential of policy promotion programs. J. Evol. Econ. 17, 299-315. http://dx.doi.org/10.1007/s00191-007-0056-6

Fritsch, M., Noseleit, M., 2013. Investigating the anatomy of the employment effect of new business formation. Camb. J. Econ. 37, 349-377. http://dx.doi.org//10.1093/cje/bes030

Gërxhani, K. 2004. The Informal Sector in Developed and Less Developed Countries: A Literature Survey. Public Choice 120:267-300. http://dx.doi.org/10.1023/B:PUCH.0000044287.88147.5e

Gil-Alana, L. A., Payne, J.E., 2015. A time-series analysis of US entrepreneurship: evidence from fractional integration. Appl. Econ. Lett. 22(7), 521-524. http://dx.doi.org/10.1080/13504851.2014.955165

Gonzalez, A., Teräsvirta, T., Van Dijk, D., Yang, Y., 2017. Panel smooth transition regression models, SSE/EFI Working Paper Series in Economics and Finance, 604. https://www.econstor.eu/bitstream/10419/56363/1/497881756.pdf

Hadri, K., 2000. Testing for Stationary in heterogeneous panel data. Economet. J. 3, 148-161. http://dx.doi.org/10.1111/1368-423X.00043

Haltiwanger, J., Jarmin, R. S., Miranda, J. 2013. Who creates jobs? Small versus Large versus Young. Rev. Econ. Stat. 95(2), 347-361- http://www.jstor.org/stable/43554390

Hansen, B.E., 1999. Threshold effects in non-dynamic panels: Estimation, testing, and inference. J. Econom. 93(2), 345-368. http://dx.doi.org/10.1016/S0304-4076(99)00025-1

Hart, D.M. (Ed), 2003. The emergence of entrepreneurship policy. Cambridge University Press, Cambridge. https://assets.cambridge.org/97805218/26778/sample/9780521826778ws.pdf

Haussen, T., Schlegel, M., 2020. Unemployment reduction through solo self-employment: A gender question?. Empir. Econ. 59(6), 3085-3105. http://dx.doi.org/10.1007/s00181-019-01754-w

Henley, A., 2021. The rise of self-employment in the UK: entrepreneurial transmission or declining job quality? Camb. J. Econ. 45(3), 457-486. http://dx.doi.org/10.1093/cje/beab007

Henley, A., 2023. Is rising self-employment associated with material deprivation in the UK? Work Employ. Soc. 37(5), 1395-1418. http://dx.doi.org/10.1177/0950017022109

Hodrick, R., Prescott, E., 1997. Postwar US Business cycles: an empirical investigation. J. Money Credit Bank 29, 1-16. http://www.jstor.org/stable/2953682

Im, K., Pesaran, M., Shin, Y., 2003. Testing for unit roots in heterogeneous panels. J. Econom. 115, 53-74. http://dx.doi.org/10.1016/S0304-4076(03)00092-7

Iversen J., Jorgensen R., Malchow-Moller, N., 2008. Defining and Measuring Entrepreneurship.

Found. Trends Entrep. 4, 1-63. http://dx.doi.org/10.1561/0300000020

Kirzner, I.M., 1973. Competition and Entrepreneurship. University of Chicago Press, Chicago.

Kremer, S., Bick A., Nautz, D., 2012. Inflation and growth: new evidence from a dynamic panel threshold analysis. Empir. Econ. 44, 861-878. http://dx.doi.org/10.1007/s00181-012-0553-9

Levin, A., Lin, C., Chu, J., 2002. Unit root tests in panel data: Asymptotic and finite-sample properties. J. Econom. 108, 1-24. http://dx.doi.org/10.1016/S0304-4076(01)00098-7

Loayza, N.V., Rigolini, J., 2011. Informal Employment: Safety Net or Growth Engine?. World Dev. 39, 1503-1515. http://dx.doi.org/10.1016/j.worlddev.2011.02.003

Lopez-Perez, E., Rodriguez-Santiago, A., and Congregado, E. 2020. Persistence in Self-Employment Rates before the Great Lockdown: The Case of the UK. Sustainability, 12(16), 6429.

Lucas, R.E.,1978. On the size distribution of business firms. Bell J. Econ. 9, 508-523. http://dx.doi.org/10.2307/3003596

Luukkonen, R., Saikkonen, P., Teräsvirta, T., 1988. Testing linearity against smooth transition autoregressive models. Biometrika 75(3), 491-499. http://dx.doi.org/10.1093/biomet/75.3.491

Maddala, G., Wu, S., 1997. On segmented multivariate regressions. Stat. Sin. 7, 497-525. https://www3.stat.sinica.edu.tw/statistica/oldpdf/A7n213.pdf

Millan, J.M., Congregado, E., Roman, C., 2014. Persistence in entrepreneurship and its implications for the European entrepreneurial promotion policy. J. Pol. Model. 36(1), 83-106. http://dx.doi.org/10.1016/j.jpolmod.2013.10.001.

Moosa, I.A., 1997. A cross-country comparison of Okun's coefficient. J. Comp. Econ. 24, 335-356. http://dx.doi.org/10.1006/jcec.1997.1433

Neumark, D., Wall, B., Zhang, J., 2011. Do Small Business Create More Jobs? New Evidence for the United States from the Nationl Establishment Time Series. Rev. Econ. Stat. 93(1), 16-29. https://doi.org/10.1162/REST_a_00060

O'Leary, D., 2022. Unemployment and entrepreneurship across high-, middle- and low-performing European regions. Reg. Stud. 9(1), 571-580. https://doi.org/10.1080/21681376.2022.2118072

Parker, S.C., 2018. The Economics of Entrepreneurship, second ed., Cambridge University Press, Cambridge UK. http://dx.doi.org/10.1017/9781316756706

Parker, S.C., Congregado, E., Golpe, A.A., 2012. Is entrepreneurship a leading or lagging indicator of the business cycle? Evidence from UK self-employment data. Int. Small Bus. J. 30(7), 736-753. https://doi.org/10.1177/0266242612437560

Pérotin, V., 2006. Entry, exit and the business cycle: are cooperatives different?. J. Comp. Econ. 34, 295-316. http://dx.doi.org/10.1016/j.jce.2006.03.002

Pfeiffer, F., Reize, F., 2000. Business start-ups by the unemployed. An econometric analysis based on firm data. Lab. Econ. 7, 629-663. http://dx.doi.org/10.1016/S0927-5371(00)00016-6

Poschke, M., 2013a. Entrepreneurs out of necessity: a snapshot. Appl. Econ. Lett. 20, 658-663. http://dx.doi.org/10.1080/13504851.2012.727968

Poschke, M., 2013b. Who becomes an entrepreneur? Labor market prospects and occupational choice. J. Econ. Dyn. Control 37, 693-710. http://dx.doi.org/10.1016/j.jedc.2012.11.003

Poschke, M., 2019. Wage Employment, Unemployment and Self-Employment across Countries. No. 12367. IZA Discussion Papers, 2019. https://docs.iza.org/dp12367.pdf

Rissman, E., 2003. Self-employment as an alternative to unemployment. Working Paper Series WP-03-34, Federal Reserve Bank of Chicago. http://dx.doi.org/10.2139/ssrn.486988

Rodriguez-Planas, N., 2010. Channels through which public employment services and small-business assistance programs work. Oxford B. Econ. Stat. 72, 458-485. http://dx.doi.org/10.1111/j.1468-0084.2010.00593.x

Rodríguez-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. Multimed. Artif. Intell. 7(3), 20-25. http://dx.doi.org/

Roman, C., Congregado, E., Millán, J.M., 2011. Dependent self-employment as a way to evade employment protection legislation. Small Bus. Econ. 37, 363-392. http://dx.doi.org/10.1007/s11187-009-9241-3

Roman, C., Congregado, E., Millán, J.M., 2013. Start-up Incentives: Entrepreneurship Policy Or Active Labour Market Programme? J. Bus. Ventur. 28, 151-175. http://dx.doi.org/10.1016/j.jbusvent.2012.01.004

Schilpzand, A., de Jong, E., 2021. Work ethic and economic development: An investigation into Weber's thesis. Eur. J. Political Econ. 66, 101958. http://dx.doi.org/10.1016/j.ejpoleco.2020.101958

Seo, M. H., and Shin, Y., 2016. Dynamic panels with threshold effect and endogeneity. J. Econom. 195(2), 169-186. http://dx.doi.org/10.1016/j.jeconom.2016.03.005

Shane, S., 2009. Why encouraging more people to become entrepreneurs is bad public policy. Small Bus. Econ. 33, 141-149. http://dx.doi.org/10.1007/s11187-009-9215-5

Thurik, A.R., Carree M.A., Van Stel A.J., Audretsch D.B., 2008. Does Self-Employment Reduce Unemployment? J. Bus. Ventur. 23, 673-686. http://dx.doi.org/10.1016/j.jbusvent.2008.01.007

Van Stel A., 2005. COMPENDIA: Harmonizing Business Ownership Data Across Countries and Over Time. Int. Entrep. Manag. J. 1(1), 105-123. http://dx.doi.org/10.1007/s11365-005-6678-y

Van Stel, A., Carree, M., Thurik, A.R.,2005. The effect of entrepreneurial activity on national economic growth. Small Bus. Econ. 24, 301-321. http://dx.doi.org/10.1007/s11187-005-1996-6

Van Stel, A., Hartog, C., Cieslik, J., 2010. Measuring Business Ownership Across Countries and Over Time: Extending the COMPENDIA Data Base. EIM Business and Policy Research, Scales Research Reports, No. H201019.

Van Stel, A., Storey, D. 2004. The Link between Firm Births and Job Creation: Is there a Upas Tree Effect? Reg. Stud. 38, 893-909. http://dx.doi.org/10.1108/JEEE-06-2020-0175

Van Stel, A., Suddle, K. 2008. The impact of new firm formation on regional development in the Netherlands. Small Bus. Econ. 30, 31-47. http://dx.doi.org/10.1007/s11187-007-9054-1

Van Stel, A., Wennekers, S., Scholman, G., 2014. Solo self-employed versus employer entrepreneurs: determinants and macro-economic effects in OECD countries. Eurasian Bus. Rev. 4, 107-136. http://dx.doi.org/10.1007/s40821-014-0003-z

Vivarelli, M., 2004. Are all the potential entrepreneurs so good? Small Bus. Econ. 23, 41-49. http://dx.doi.org/10.1023/B:SBEJ.0000026023.11752.a9

Wennekers, S., Thurik, A.R., 1999. Linking entrepreneurship and economic growth. Small Bus. Econ. 13, 27-55. http://dx.doi.org/10.1023/A:1008063200484

Wennekers, S., van Stel, A., Thurik, R., Reynolds, P. , 2005. Nascent Entrepreneurship and the Level of Economic Development. Small Bus. Econ. 24, 293-309. http://dx.doi.org/10.1007/s11187-005-1994-8

Wong, P.K., Ho, Y.P., Autio E. 2005. Entrepreneurship, Innovation and Economics Growth. Evidence from GEM data. Small Bus. Econ. 24, 335-350. http://dx.doi.org/10.1007/s11187-005-2000-1