Evaluating the long-term impacts of economic or policy shocks among necessity and opportunity entrepreneurs

E. Congregado, D. Troncoso, N. Rubino

October 2022

Abstract

We aim at distinguishing between two categories of self-employed workers based on revealed intentions (necessity versus opportunity entrepreneurs), equating hysteresis with the existence of a unit root in a variable whether aggregate rates of entrepreneurship exhibit persistence or hysteresis, and finally use longer time horizons in formal evaluation exercises rather than the few years which are commonly used to gauge entrepreneurship policy impacts. Our analysis includes a timewise analysis of presistence, checks for nonlinear patterns and finally models regime switching behavior and markov-based probability persistence analysis for the aforementioned groups.

- JEL: C32 E23 J24 M13
- Keywords: necessity entrepreneurs; transition regression model; hysteresis and stochastic components: nonlinearities; unit roots

Contents

1 Introduction

The dynamics of the self-employment rate during the business cycle keep being a source of controversy among scholars, summarized in the so-called push and pull hypotheses [1], as well as in the distinction between opportunity and necessity entrepreneurs [2], as two different components of business creation with potentially opposite dynamics over the business cycle. Empirical estimates of the self-employment/unemployment relationship only aspired to capture a “net” effect of the recession-push and the prosperity-pull effects [3]. However, recent literature has provided operational definitions of opportunity and necessity entrepreneurship using readily available nationally representative data [4]-[5], applying time series techniques for checking the macro-dynamics of opportunity and necessity self-employment during the business cycle [6]-[7]. With different targets –contributions to innovation and economic growth or as an alternative of other ALMP’s– governments, around the world, have devised and implemented portfolios of policies to promote entrepreneurship (more appropriate opportunity entrepreneurs) or to turn unemployment into self-employment (more appropriate for necessity motivated entrepreneurs). These interventions impose sizeable costs on the taxpayer. For this reason, these policies (in both the short and long term) should be properly monitored and evaluated. Recent studies have evaluated national experiences employing register data by using different outcomes such as duration into self or regular employment, incomes or job satisfaction [8]. In general, this literature appears to point at the positive results of start-up programs for unemployed and underline higher probabilities of survival among opportunity entrepreneurs. [9]. However, and due to data limitations, the long-run effects are not accurately captured by conventional evaluations, which are usually performed a few years after the policies are implemented, and so capture only short-term impacts. There exists, however, an alternative path: exploring the macro-dynamics of self-employment, in particular whether entrepreneurship evolves as a trend stationary or as a non-stationary time-series process. If entrepreneurship is trend stationary, economic and policy shocks can be regarded as transitory from an aggregate perspective: the rate of entrepreneurship eventually reverts to its underlying, long-run rate. If, on the other hand, the rate of entrepreneurship is non-stationary, such shocks will have permanent effects. With this in mind, and to the best of our knowledge this paper represents the first attempt to evaluate the long run effects of positive shocks in the self-employment due to push factors, at the macro-level. Spain is a suitable case of study since e.g. the use of entrepreneurship promotion as an active labor market policy has been intensively applied as a way to combat unemployment. By using a new data set of regional time series of necessity entrepreneurs for the 17 Spanish regions, our results points to the existence of hysteresis. As the main analysis entails unit root testing in a time series environment, we decided to employ panel unit root testing with alternative heterogeneous hypotheses, robust to both classical spherical disturbances and eventual spatial spill overs, in order to check for the robustness of the main results. Our analysis adds to the national and regional level estimates an additional layer of complexity by estimating nonlinear models of convergence in five different economic sectors, agriculture, industry, low and high skilled services. As a final exercise, we run a Markov swtiching autoregressive model to test for the persistence of the probability across low and high states of numerosity of the aforementioned necessity and opportunity entrepreneur groups.

Our analysis shows that both groups most likely follow an integrated process; that variables such as the rate of unemployment and GDP (proxyed by an industrial production index) perform well in defining a nonlinear regime switching model which underlines, under given conditions, the cyclical behavior of the opportunity and necessity entrepreneur groups, with the proxyed GDP in particular describing a nonlinear behavior in any given Spanish region when the opportunity group is considered; that the probability of maintaining a given numerosity level for both groups is closer to unity than to 0, indicating that economic operators react with some delay to economic news and are perhaps less reactive than we would expect to exogenous shocks such the recent financial crisis, but nevertheless showing that the necessity group is more likely to transition back to a lower numerosity given any possible exogenous change from period to period, while the opportunity group stands out as the more resilient one.

2 Data

All Data, besides the CVEC (Data corrected for seasonal and specific variations) GDP index from AIREF (Autoridad Indipendiente de Responsabilidad Fiscal), are drawn from the quarterly microdata 2000/1 to 2020/4 of the Spanish

Labor Force Survey (EPA). The Survey, conducted by the National Statistical Institute (INE), is a large household sample survey providing results on labor participation of people aged 16 and over as well as persons outside the labor force in which each sampled individual remaining in the survey for a period of six quarters at a time, with no resampling after individuals are rotated out of the sample.

The Survey is targeted at a rotating sample of around 60,000 households throughout the national territory. For every household member, both socioeconomic and labor information is collected in order to summarize the main characteristics of the Spanish workforce each quarter. Individuals in the sample are interviewed for six consecutive quarters, thus we have information on quarterly labor transitions for a maximum period of 18 months for each individual in the sample.

Following Fairlie and Fossen [4]-[5] we will define necessity entrepreneurs as self-employed who answered their previous labor force status question as "out of the labor force but able to work". In contrast, opportunity entrepreneurs were those individuals whose status was 'in the labor force' and they had been "employed" for either a public or private institution preceding their current self-employment status.

By using this approach once the two categories of self-employed workers are identified, and applying the elevation factors, we built the national aggregate quarterly time series of necessity and opportunity entrepreneurs and a panel data for the 17 Spanish regions.

After applying the necessary filters to the data, our panel contains a total of 11,849,287 individual observations. For each of the above mentioned groups, each time series has been disaggregated by Spanish regions and, for each region, by: 1) the duration of the unemployment period prior to the transition to self-employment (only for necessity entrepreneurs); 2) the educational attainment level; and 3) the economic sector observed in self-employment transition.

We define a necessity entrepreneur in quarter t as an individual who experiences a transition towards self-employment state between quarters t and $t+1$, conditional on being observed in quarter t in a non-employment state (either by being classified as unemployed -actively seeking employment-, or as inactive). In contrast, we define an opportunity entrepreneur in quarter t as an individual who experiences a transition towards self-employment state between quarter t and $t+1$, conditional on being observed in quarter t in an employment status other than self-employment (i.e. wage employees, both in the public and private sector).

Therefore, the total number of individuals transiting to self-employment state between quarter t and t+1 is composed of those come from non-employment state (necessity entrepreneurs) and those who come from employment state (opportunity entrepreneurs). Based on this, we analyze the share of each of these two types of entrepreneurs over the total at each quarter.

3 Regarding the test equation

Our analysis entails a unit root and stationarity testing of the raw variables. Through a series of Dickey Fuller based tests and considering both a linear and a nonlinear alternative [10]-[11], we test for the presence of persistent behavior (stochastic trends) in a series of quarterly observations spanning the 2000 - 2020 period. As a second step, we make use of the Hansen test [12], in its panel rendition, to check for possible alternatives to non linearity by comparing a sequence of F tests based on the null hypothesis of non linearity against the possible alternative of possible existing branches, namely up to two. We then test for possible discrete regime switching behavior –Threshold autoregression models (TAR) and self exciting threshold autoregression models (SETAR) [13] – by selecting an array of possible weakly exogenous threshold variables (unemployment, changes in unemployment, GDP levels and changes proxyed by an industrial production index) and choosing the favored model based on Information Criteria minimization against the linear alternative. Robustness of the unit root analysis is checked through Panel unit root tests which allow for heterogeneous hypothesis testing (relaxing the hypothesis of a single coefficient across all cross sections [14]-[15]), while possible asymmetries are checked via the aforementioned Hansen Panel test as in Terasvirta et al. [16]. As we have stated in the introduction the analysis is completed with the estimation of an autoregressive Markov switching time series model, which gave us further insights on the persistence of the series. Further details on the Markov switching application methods are contained in the related section.

Our test equation, across the whole analysis, is an ADF type test equation with an attractor to fir the SETAR modelling in order to test for different speed of return to the unconditional mean.

$$\Delta nec_t = c + \rho nec_{t-1} + e_t \tag{1}$$

As the hysteresis hypothesis represents a crucial component of our work, we have chosen to check for the robustness of the one and two breaks unit root models by employing an additional unit root test with some additional properties. To be more specific, we choose to select a second generation panel unit root test, capable of taking into account possible cross sectional correlation, to countercheck the results of time series testing. Our choice ultimately fell on the Im, Pesaran and Shin unit root test and Hadri's stationarity test. The latter was chosen as it presents an inverted null of stationarity, takes possible correlation in the error terms of the cross sections into account and even more importantly present an heterogeneous alternative instead of a standard homogeneous one (meaning that the null hypothesis implies all cross sections are stationary while the alternative states that some cross sections contain unit roots). We reckon such choices and strategies would end up being adequate to our set up (17 different autonomous regions with different economic structures), considering homogeneity across regions members of the country on top of time variation.

4 Unit root Analysis

In our paper, we employ both additive and innovation model specifications from Perron and Vogelsang unit root break test. The main difference between the two specifications is that the innovation model entails a less sudden change in the level relationship, while the additive model, which we would focus on given the results of visual inspection, would present an alternative based on a sharper change of the intercept.. The test equation below (2) would thus represent a unit root testing with a nested alternative of a break in the constant.

$$nec_t = \mu + \theta DU_t(T_b) + \alpha nec_{t-1} + \Sigma_{j=1}^{k} c_j \Delta nec_{t-j} + w_t \tag{2}$$

The Perron-Vogelsang is perhaps a common test in literature. In order to accommodate a systematic shift in our data and to avoid data mining[1], we choose to extend the analysis with an additional test, which would add a second deterministic break in levels. Such test is basically a generalization from Clemente-Montañes-Reyes of the already cited Perron-Vogelsang test, and its test equation is visible in (eq-3).

$$\begin{aligned} nec_t \;=\; & \mu + \theta_1 DU_{1,t}(T_{a,t}) + \theta_2 DU_{2,t}(T_{b,t+s}) + \alpha nec_{t-1} + \\ & \Sigma_{j=1}^{k} c_j \Delta nec_{t-j} + w_t \end{aligned} \tag{3}$$

In both Equations 2 and 3, the homogeneous hypothesis will thus be H_0 : $\alpha = 0$ against a heterogeneous null $H_a = \alpha < 0$.

4.1 Perron-Vogelsang results with a single break

For the remainder of the paper, we shall use the following numbering to indicate the autonomous communities of Spain and the country as a whole: 1, Andalucia; 2, Aragon; 3, Asturias; 4, Balearic Islands; 5, Canary Islands; 6, Cantabria; 7, Castile and Leon; 8, Castile and La Mancha; 9, Catalonia; 10, Valencian Community; 11, Extremadura; 12, Galicia; 13, Community of Madrid; 14, Murcia; 15, Navarre; 16, Basque Country; 17, La Rioja; 18, Spain as a whole.

Let's consider the innovation outlier first: for the necessity entrepreneurs variable, the hypothesis of unit root with one break could not be rejected at 5% in Cantabria, Castile-la Mancha, Valencia, Galicia and Madrid (6 8 10 12 13) while it was generally accepted elsewhere in the country. As for the opportunity entrepreneurs variable, the null of nonstationarity was found to be generally unrejected in Andalucia, Asturias, Cantabria, Murcia, Navarre and the Basque Country (1 3 6 14 15 16). Unemployment was never found to be stationary and break dates for the alternatives would be closely contained between the 2007q3 and the 2008q3 interval.

[1] On such issue, please have a look at Figure 1 to Figure 4. The data we could retrieve from EPA allowed us to build series with a systematic level change in 2004. In order to check for consistency of results across same unit root testing applications and considering the somewhat limited length of our time series we necesserely had to resort to up to two break alternatives.

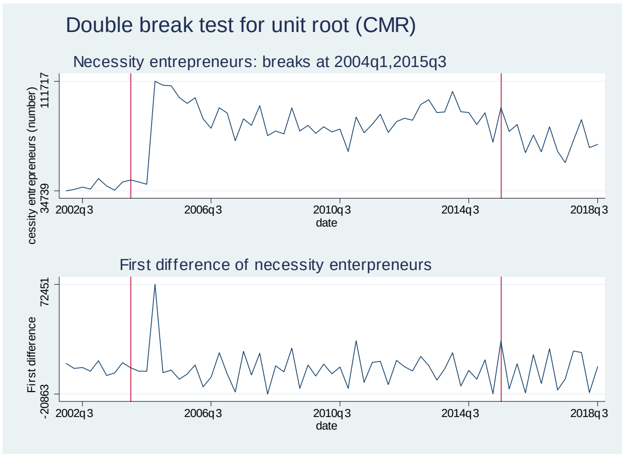

Figure 1: Unit Root test, two breaks, necessity entrepreneurs

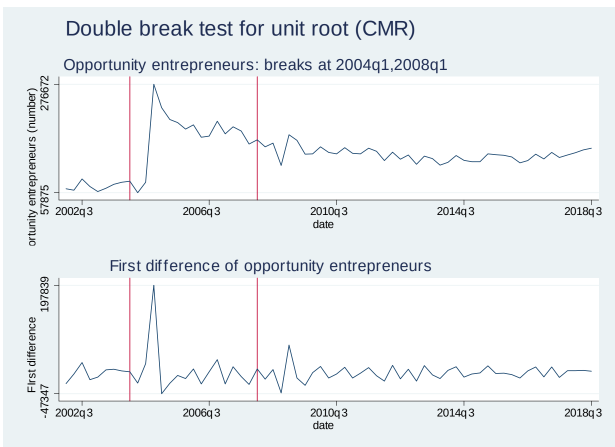

Figure 2: Unit Root test, two breaks, opportunity entrepreneurs

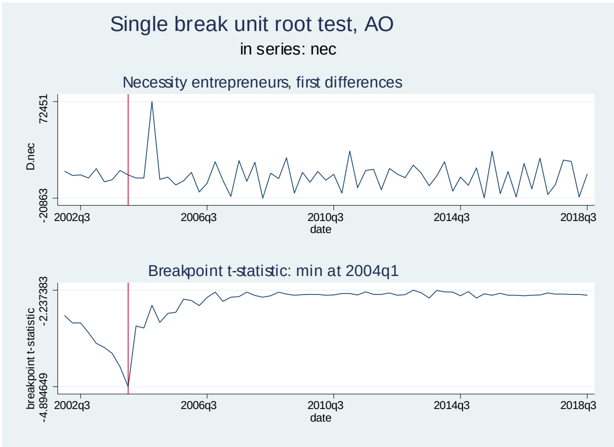

Figure 3: Unit root test, one break, AO, necessity workers

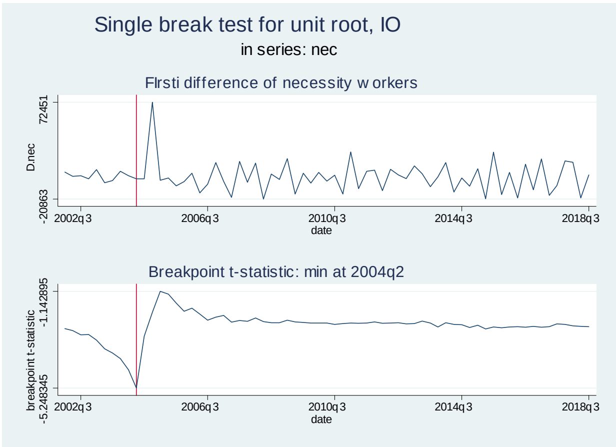

Figure 4: Unit root test, one break, IO, necessity workers

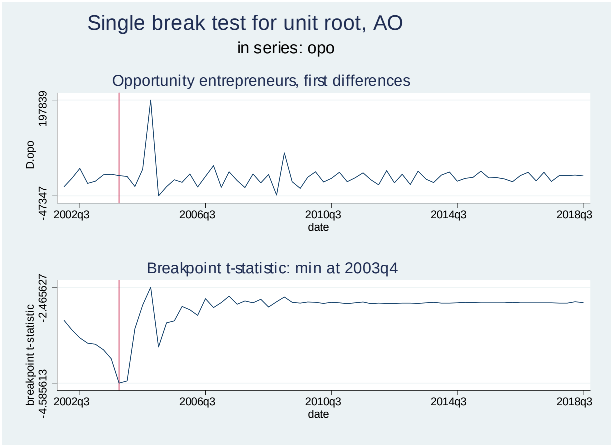

Figure 5: Unit root test, one break, AO, opportunity entrepreneurs

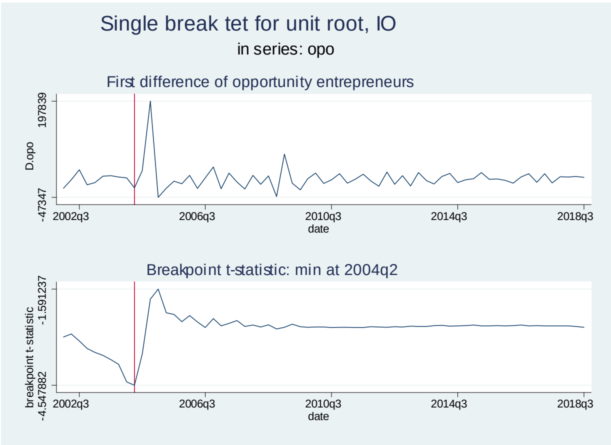

Figure 6: Unit root test, one break, IO, opportunity workers

Table 1: Unit Root tests, AO, one break, Necessity entrepreneurs

Regions	*t-stat*	*Break 1 stat*	*Break 2 stat*	*5% crit*
(I)	(II)	(III)	(IV)	(V)
Andalusia	-4,968*	4,129		-4,270
Aragon	-8,093*	3,715		-4,270
Asturias	-7,629*	3,609		-4,270
Balearic Islands	-4,859*	3,881		-4,270
Canary Islands	-5,374*	4,134		-4,270
Cantabria	-3,729	1,520		-4,270
Castile and Leon	-8,224*	4,163		-4,270
Castile-La Mancha	-3,325	2,512		-4,270
Catalonia	-8,237*	5,894		-4,270
Valencian Community	-3,844	3,107		-4,270
Extremadura	-8,472*	3,485		-4,270
Galicia	-1,824	-2,602		-4,270
Community of Madrid	-4,132	3,411		-4,270
Region of Murcia	-10,370*	4,862		-4,270
Navarre	-7,330*	-1,475		-4,270
Basque Country	-4,506*	3,467		-4,270
La Rioja	-9,126*	3,784		-4,270
SPAIN	-5,579*	4,599		-4,270

Unit root tests. t-statistics of the endogenously retrieved breaks reported. One break case.

Table 2: Unit Root tests, AO, one break, Necessity entrepreneurs

Regions	*t-stat*	*Break 1 stat*	*Break 2 stat*	*5% crit*
(I)	(II)	(III)	(IV)	(V)
Andalucia	-3,388	8,850		-3,560
Aragon	-8,040*	3,814		-3,560
Asturias	-1,793	4,544		-3,560
Balearic Islands	-4,448*	4,856		-3,560
Canary Islands	-6,338*	5,477		-3,560
Cantabria	-3,248	2,919		-3,560
Castile and Leon	-0,376	4,242		-3,560
Castile-La Mancha	-2,471	5,258		-3,560
Catalonia	-4,531*	8,703		-3,560
Valencian Community	-7,948*	6,289		-3,560
Extremadura	-3,440	3,850		-3,560
Galicia	-1,252	2,990		-3,560
Community of Madrid	-1,170	7,893		-3,560
Region of Murcia	-7,159*	4,005		-3,560
Navarre	-8,042*	2,841		-3,560
Basque Country	-2,944	5,045		-3,560
La Rioja	-9,230*	3,984		-3,560
SPAIN	-2,269	11,273		-3,560

Unit root tests. t-statistics of the endogenously retrieved breaks reported. One break case.

The additive outlier, on the other side, gave us the following results: in the case of necessity entrepreneurs, Andalucia, Asturias, Cantabria, Castile and Leon, Castile-La Mancha, Extremadura, Galicia, Madrid, the Basque Country and the whole country series (1 3 6 7 8 11 12 13 16 18) could not see the null of unit root rejected for an alternative with a sudden change structure. As for opportunity entrepreneurs, the majority of the regions, Andalucia, Aragon, Asturias, Balearic Islands, Castile and Leon, Castile-La Mancha, Catalonia, Valencia, Extremadura, Madrid, Murcia, Navarre, Basque Country and the whole country series (1 2 3 4 7 8 9 10 11 13 14 15 16 18) were considered I(1) by the test. As with the IO model, unemployment appears to be absolutely nonstationary at 5% in any given region.

4.2 Clemente-Montañes-Reyes with two breaks

Let us turn our attention to the 2 breaks model: In the case of the necessity entrepreneurs, the null hypothesis is not rejected for Asturias, Balearic islands, Cantabria, Castile-La Mancha, Valencia, Galicia and the Basque Countries (3 4 6 8 10 12 16) with the innovation outlier model at 5%. As for the more sudden

Table 3: Unit Root tests, IO, one break, Opportunity entrepreneurs

Regions	*t-stat*	*Break 1 stat*	*Break 2 stat*	*5% crit*
(I)	(II)	(III)	(IV)	(V)
Andalucia	-3,280	-1,763		-4,270
Aragon	-6,199*	3,640		-4,270
Asturias	-4,101	2,721		-4,270
Balearic Islands	-8.527*	4,447		-4,270
Canary Islands	-4,476*	2,333		-4,270
Cantabria	-3,315	1,343		-4,270
Castile and Leon	-5,327*	2,926		-4,270
Castile-La Mancha	-5,554*	3,535		-4,270
Catalonia	-8,099	6,177		-4,270
Valencian Community	-6,515*	4,511		-4,270
Extremadura	-5,620*	2,429		-4,270
Galicia	-4,536*	2,817		-4,270
Community of Madrid	-7,056*	6,658		-4,270
Region of Murcia	-3,476	-3,319		-4,270
Navarre	-3,779	2,495		-4,270
Basque Country	-3,487	2,562		-4,270
La Rioja	-8,473*	5,603		-4,270
SPAIN	-5,311*	4,020		-4,270

Unit root tests. t-statistics of the endogenously retrieved breaks reported. One break case.

Table 4: Unit Root tests, AO, one break, Opportunity entrepreneurs

Regions	t-stat	Break 1 stat	Break 2 stat	5% crit
Andalucia	-3,190	4,872		-3,560
Aragon	-1,051	5,962		-3,560
Asturias	-2,374	3,601		-3,560
Balearic Islands	-2,641	5,288		-3,560
Canary Islands	-4,874*	4,364		-3,560
Cantabria	-6,415*	1,751		-3,560
Castile and Leon	-1,801	5,421		-3,560
Castile-La Mancha	-2,199	6,447		-3,560
Catalonia	-3,399	9,367		-3,560
Valencian Community	-2,586	7,448		-3,560
Extremadura	-2,503	4,136		-3,560
Galicia	-4,603*	2,550		-3,560
Community of Madrid	-1,950	7,677		-3,560
Region of Murcia	-1,998	-2,732		-3,560
Navarre	-3,297	3,800		-3,560
Basque Country	-3,014	5,880		-3,560
La Rioja	-5,769*	7,660		-3,560
SPAIN	-2,241	8,158		-3,560

Unit root tests. t-statistics of the endogenously retrieved breaks reported. One break case.

Table 5: Unit Root tests, IO, two breaks, Necessity entrepreneurs

Regions	*t-stat*	*Break 1 stat*	*Break 2 stat*	*5% crit*
(I)	(II)	(III)	(IV)	(V)
Andalusia	-9,486*	7,433*	-4,742	-5,490
Aragon	-8,300*	-3,657	-0,743	-5,490
Asturias	-4,451	3,687	-1,244	-5,490
Balearic Islands	-5,094	3,985	1,467	-5,490
Canary Islands	-9,375*	5,639	-3,935	-5,490
Cantabria	-3,407	2,735	-1,979	-5,490
Castile and Leon	-10,188*	6,052	-4,090	-5,490
Castile-La Mancha	-3,850	3,204	-1,879	-5,490
Catalonia	-7,211*	6,254	-3,647	-5,490
Valencian Community	-4,901	4,295	-3,032	-5,490
Extremadura	-9,474*	8,051	-6,824	-5,490
Galicia	-4,368	5,291	-6,187	-5,490
Community of Madrid	-9,969*	7,476	-4,366	-5,490
Region of Murcia	-11,136*	6,135	-3,121	-5,490
Navarre	-8,007*	3,008	-3,352	-5,490
Basque Country	-4,502	4,918	-3,224	-5,490
La Rioja	-9,311*	3,224	0,688	-5,490
SPAIN	-8,991*	8,053	-5,888	-5,490

Unit root tests. t-statistics of the endogenously retrieved breaks reported. Two breaks case.

change implied by the additive outlier model, again in the case of necessity entrepreneurs, the null hypothesis is not rejected for Aragon, Balearic Islands, Canary Islands, Castile and Leon, Castile-La Mancha, Catalonia, Extremadura, Galicia, Madrid, Murcia, Basque Country and La Rioja (2 4 5 7 8 9 11 12 13 14 16 18) with the additive outlier model at 5%. Given the sudden sharp change we could detect by the means of a preliminary visual inspection, we reckon the last model would be indeed more representative of the true process underlying the series for the necessity entrepreneurs group. In the case of the opportunity entrepreneurs, the innovative outlier test fails to reject the unit root against two deterministic jumps in the mean in Navarre only, while the additive outlier test states that regions Aragon, Asturias, Cantabria, Castile and Leon, Castile-La Mancha, Catalonia, Valencia, Galicia, Murcia, Navarre and the Basque Country (2 3 6 7 8 9 10 12 14 15 16 18) could not reject the unit root against the deterministic hypothesis at the 5% percent significance level. In general, the endogenously retrieved break dates point at the first quarter of 2004 and at the neighborhood of 2008 as the suggested jump dates. As the first date presents a generally sharper shift in the relationship and is more closely connected to the nature of the data, the second one would most likely connect to the financial crisis period, thus disconnecting the causes of the long memory process from the exogenous impact of the sub-prime and subsequent economic crisis in the European Area. Unsurprisingly, the unemployment rate was found to be nonstationary in almost every region, with the retrieved break dates pointing at 2008 and 2015 respectively.

4.3 Results of the hysteresis analysis at the national level

National level results are merged and showed jointly in Table 9. As we have stated previously, the innovative outlier model easily rejects the unit root hypothesis with flying colors. However, considering how sudden changes in employment would appear by visual inspection and considering the width of the series, the additive outlier specification appears to be unable to do so in both

Table 6: Unit Root tests, AO, two breaks, Necessity entrepreneurs

Regions	*t-stat*	*Break 1 stat*	*Break 2 stat*	*5% crit*
(I)	(II)	(III)	(IV)	(V)
Andalucia	-8,442*	8,095	-2,704	-5,490
Aragon	-2,774	2,961	0,921	-5,490
Asturias	-6,954*	4,787	-1,786	-5,490
Balearic Islands	-3,403	4,023	1,331	-5,490
Canary Islands	-4,596	6,957	-4,106	-5,490
Cantabria	-10,326*	3,337	-1,708	-5,490
Castile and Leon	-4,668	5,050	-2,996	-5,490
Castile-La Mancha	-3,120	5,023	-0,528	-5,490
Catalonia	-5,184	9,439	-2,730	-5,490
Valencian Community	-8,817*	7,005	-2,683	-5,490
Extremadura	-3,565	5,497	-3,645	-5,490
Galicia	-1,635	6,425	-5,315	-5,490
Community of Madrid	-5,132	9,288	-3,899	-5,490
Region of Murcia	-2,503	5,673	-2,795	-5,490
Navarre	-6,589*	3,279	-1,765	-5,490
Basque Country	-1,492	5,594	-2,795	-5,490
La Rioja	-9,246*	3,042	1,001	-5,490
SPAIN	-4,380	13,533	-4,773	-5,490

Unit root tests. t-statistics of the endogenously retrieved breaks reported. Two breaks case.

Table 7: Unit Root tests, IO, two breaks, Opportunity entrepreneurs

Regions	*t-stat*	*Break 1 stat*	*Break 2 stat*	*5% crit*
(I)	(II)	(III)	(IV)	(V)
Andalucia	-8,282*	8,203	-8,457	-5,490
Aragon	-7,579*	6,861	-6,284	-5,490
Asturias	-10,409*	7,748	-6,118	-5,490
Balearic Islands	-9,407*	5,558	-3,049	-5,490
Canary Islands	-8,228*	4,992	-3,811	-5,490
Cantabria	-10,020*	7,145	-6,495	-5,490
Castile and Leon	'-10,059*	8,859	-7,968	-5,490
Castile-La Mancha	-6,784*	-6,892	-7,180	-5,490
Catalonia	-10,550*	8,768	-5,190	-5,490
Valencian Community	-7,751*	9,250	-7,285	-5,490
Extremadura	-10,194*	7,665	-7,206	-5,490
Galicia	-7,435*	8,160	-7,127	-5,490
Community of Madrid	-8,572*	8,949	-5,746	-5,490
Region of Murcia	-9,447*	5,454	-5,894	-5,490
Navarre	-5,180	6,364	-5,937	-5,490
Basque Country	-7,662*	7,600	-7,091	-5,490
La Rioja	-10,090*	7,366	-4,398	-5,490
SPAIN	-12,664*	11,930	-10,383	-5,490

Unit root tests. t-statistics of the endogenously retrieved breaks reported. Two breaks case.

Table 8: Unit Root tests, AO, two breaks, Opportunity entrepreneurs

Regions	*t-stat*	*Break 1 stat*	*Break 2 stat*	*5% crit*
Andalucia	-8,225*	9,855	-7,719	-5,490
Aragon	-2,222	8,558	-5,179	-5,490
Asturias	-2,492	8,336	-4,882	-5,490
Balearic Islands	-7,681*	5,964	-2,432	-5,490
Canary Islands	-8,130*	5,217	-2,883	-5,490
Cantabria	-3,488	5,596	-4,449	-5,490
Castile and Leon	-3,195	9,504	-6,353	-5,490
Castile-La Mancha	-4,894	9,139	-5,298	-5,490
Catalonia	-4,110	10,037	-3,135	-5,490
Valencian Community	-5,393	9,143	-4,437	-5,490
Extremadura	-7,692*	7,723	-6,094	-5,490
Galicia	-3,482	7,939	-5,563	-5,490
Community of Madrid	-8,434*	8,113	-2,671	-5,490
Region of Murcia	-3,145	5,569	-5,810	-5,490
Navarre	-1,909	5,066	-3,145	-5,490
Basque Country	-3,486	8,103	-4,714	-5,490
La Rioja	-6,919*	7,528	-3,372	-5,490
SPAIN	-2,948	12,076	-6,261	-5,490

Unit root tests. t-statistics of the endogenously retrieved breaks reported. Two breaks case.

Table 9: Unit Root tests, National level

	t	*Break 2 p-value*	*Break 2 p-value*	*5 % critical value*
(I)	(II)	(III)	(IV)	(V)
nec_t				
AO 1 break	-2,27	0,00		-3,56
IO 1 break	-5,58*	0,00		-4,27
AO 2 breaks	-4,38	0,00	0,00	-5,49
IO 2 breaks	-9,00*	0,00	0,00	-5,49
opo_t				
AO 1 break	-2,24	0,00		-3,56
IO 1 break	-5,31*	0,00		-4,27
AO 2 breaks	- 2,94	0,00	0,00	-5,49
IO 2 breaks	-12,66*	0,00	0,00	-5,49

Unit root tests. P values of the endogenously retrieved breaks reported. One and two breaks hypothesis.

one and two break model tests. In order to make up our mind and to be able to take a final decision on the overall hysteresis phenomenon in the country as a whole, we would have to resort to a counter-test to act as a robustness check. That is exactly what we shall see in the section dedicated to the panel unit root tests.

4.4 Sectorial analysis at the autonomous regional level

We finally test for unit root behavior at the sectorial level, to provide some understanding on how the hysteresis phenomenon behaves across Agriculture, Industry, Construction, Low Qualified Services and High Qualified Services.

Given the initial results of the unit root analysis for the regions as a whole, we now look at the sectorial evolution of the memory process for the necessity entrepreneurs group focusing in particular on the additive outlier model. That appears considering the outcome of our previous inspections (visual and stability analysis) to be the choice most suited to the nature of our data.

Starting with the necessity entrepreneurs group and in the first sector, agriculture, we could not reject the unit root null for the level variables in Aragon, Castile and Leon, Castile and La Mancha, Catalonia, Extremadura, Galicia, Navarre and the country as a whole (2 7 8 9 11 12 15 18). As for the second sector, the industrial one, unit root behavior was found in the Balearic Islands, Navarre, Basque Country and LA Rioja (4 15 16 17). As we switch our attention to the construction sector, testing for a unit root shows that regions Canary Islands, Castile and La Mancha, Valencian Community and Spain as a whole (5 8 10 18) would have indeed a unit root hidden in their data generating process. Let us complete the analysis of the necessity entrepreneurs group with the service sector, as we can distinguish between low skilled and high skilled entrepreneurs in this instance. For the share of the service sector requiring less qualified labor, a unit root was found in Andalucia, Aragon, Asturias, Canary Islands, Valencian Community, Galicia, Community of Madrid, Murcia and Spain as a whole (1 2 3 5 10 12 13 14 18). Finally, high skilled labor led to non rejection of the unit root hypothesis in Galicia and Navarre only (12 15).

Let us focus on the opportunity entrepreneurs group now. In the agricultural sector, regions Asturia, Castile and Leon, Catalonia, Extremadura, Navarre, La

Rioja and Spain as a whole (3 7 9 11 15 17 18) showed evidence of unit root behavior. As for the industrial sector, we confirmed a unit root in Aragon, Asturias, Canary Islands, Castile and La Mancha, Extremadura, Galicia, Community of Madrid, Basque Country and Spain as a whole (2 3 5 8 11 12 13 16 18). In the construction sector, Aragon, Asturias, Balearic Islands, Castile and la Mancha, Galicia and Spain as a whole (2 3 4 8 12 18) were found to show a unit root behavior. Finally, in the low and high skill service sectors, regions Andalucia, Aragon, Asturias, Balearic Islands, Castile and Leon, Extremadura, Community of Madrid, Navarre, Basque Country and Spain as a whole (1 2 3 4 7 11 13 15 16 18) presented a unit root for the former sector and Canary Islands and Basque Country (5 16) only for the latter. What indeed appears striking is the strong degree of persistence found in low skilled labor sectors, which is in relative terms much less present in the high skilled sector.

As a single break alternative naturally suits our data by construction, it is crucial for us to add and compare the results illustrated in the last paragraph with a two deterministic breaks alternative. We thus employ once more the Clemente et al. extending the previous analysis once more to all available in regions across all sectors with a sharp change as implied by the additive outlier model.

As before, let us focus once again on the necessity entrepreneurs first. For the agricultural sector, Andalucia, Balearic Islands, Cantabria, Castile and Leon, Castile and La Mancha, Catalonia, Valencian Community, Extremadura, Galicia, Navarre and Spain as a whole (1 4 6 7 8 9 10 11 12 15 18) showed a unit root process with probability values frequently lower than one percent. For the industry sector, the series for Andalucia, Asturias, Balearic Islands, Canary Islands, Cantabria, Valencian Community, Extremadura, Galicia, Murcia, Navarre, La Rioja (1 3 4 5 6 10 11 12 14 15 17) were seen as integrated by the test. In the case of the construction sector, series for Aragon, Canary Islands, Castile and Leon, Castile and La Mancha, Catalonia, Valencian Community, Extremadura, Navarre La Rioja and Spain as a whole (2 5 7 8 9 10 11 15 17 18) contained a unit root. Finally, for the low skilled service sector and the high skilled service sector, series for Asturias, Balearic Islands, Canary Islands, Valencian Community, Navarre, La Rioja and Spain as a whole (3 4 5 10 15 17 18) appear to be integrated of order one for the former, while series for Asturias, Cantabria, Castile and La Mancha, Extremadura, Murcia and Navarre) 3 6 8 11 14 15 appear to be integrated of order one for the latter.

Let's finally turn our attention to the opportunity entrepreneurs series: in the agricultural sector, Aragon, Balearic Islands, Cantabria, Castile and Leon, Castile and La Mancha, Catalonia, Valencian Community, Extremadura, Galicia, Community of Madrid, Navarre, Basque Country, La Rioja and Spain as a whole (2 4 6 7 8 9 10 11 12 13 15 16 17 18) presented a unit root. In the industrial sector, a non stationary process could be confirmed for Aragon, Asturias, Castile and Leon, Catalonia, Community of Madrid, Murcia, Basque Country, La Rioja and the country as a whole (2 3 7 9 13 14 16 17 18). Switching to construction, we find a unit root process in the opportunity entrepreneurs series for Andalucia, Aragon, Balearic Island, Castile and La Mancha, Catalonia, Valencian Community, Galicia, Murcia, La Rioja and Spain as a whole (1 2 4 8 9 10 12 14 17 18). Turning our view to the service sector, in the case of the low qualified services, we have decisive evidence of a unit root process in Aragon, Asturias, Balearic Islands, Cantabria, Castile and Leon, Catalonia, Valencian Community, Galicia, Navarre, Basque Country and the country as a whole (2 3 4 6 7 9 10 11 12 15 16 18). Finally, for the higher skilled labour force in the service sector, we have evidence of a unit root process in the Balearic Islands, Canary Islands, Cantabria, Castile and Leon, Extremadura, Galicia, Community of Madrid, Basque Country, La Rioja and Spain as a whole (4 5 6 7 11 12 13 16 17 18). **All in all, if not for the high degree of the heterogeneity of the results on a region by region basis, this paragraph estimates prove that persistence, as an overall country phenomenon, is indeed present across almost all sectors when national series are considered.**

4.5 Robustness analysis - Panel Unit Root

As we have stated previously, we have chosen Panel unit root tests to countercheck for the results of the stochastic time series persistence analysis on the basis of an inverted null, accountability of cross sectional dependence (and of course common spherical disturbances), and heterogeneity of the alternative hypothesis. We remind that, similarly to the time series case, the ADF model test equation for the Panel unit root from Im, Pesaran and Shin (with a null of non stationarity and omitting panel specific trends) is:

$$\Delta nec_{i,t} = \xi_i nec_{i,t-1} + \epsilon_{i,t} \tag{4}$$

so that the homogeneous hypothesis will be $H_0 : \xi_i = 0$ versus a heterogeneous null $H_a = \xi_j < 0$, where $j \in [1 : i]$ On the contrary, the LM statistic for the Hadri test, which presents a null of stationarity, is modelled as an LM test of the variances of a random walk with trend model and the random walk itself. Considering the model:

$$nec_{i,t} = w_{i,t} + \beta_i t + \epsilon_{i,t} \tag{5}$$

with random walk defined as:

$$w_{i,t} = w_{i,t-1} + \varepsilon_{i,t} \tag{6}$$

the test could be written as: $H_0 : \vartheta = \frac{var(\varepsilon_{i,t})}{var(\epsilon_{i,t})} = 0$ versus an alternative $H_a : \vartheta > 0$. Given the random walk collapses to a constant if its variance is 0, the model would become almost deterministic with an i.i.d. stochastic component, and the series would be automatically trend stationary. Logically, the null implies stationarity, not unit root behavior

The results of Hadri's second generation unit test for the whole country are thus available in Table 10, as well as the results of the Im, Pesaran and Shin first generation unit root test. The outcome of the analysis strongly rejects the unified null of stationarity and opts to point out at the alternative, heterogeneous possibility that some units (cross sections) might indeed be integrated.

Table 10: Unit Root tests, homogeneous alternative, CSC and heteroskedasticity, Panel

	Test	H_0	H_a	*Heteroskedasticity*	*CSS*	*Statistic*	*p-value*
(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(VIII)
Δopo_t	Hadri	Hom. Stationarity	Het. Unit Root	Robust	Robust	-4,10	1,00
opo_t	Hadri	Hom. Stationarity	Het. Unit Root	Robust	Robust	16,10*	0,00
Δnec_t	Hadri	Hom. Stationarity	Het. Unit Root	Robust	Robust	-4,22	0,00
nec_t	Hadri	Hom. Stationarity	Het. Unit Root	Robust	Robust	22,26*	0,00
Δopo_t	IPS	Hom. Unit Root	Het. Stationarity	Robust	Demeaned	-31,44*	0.00
opo_t	IPS	Hom. Unit Root	Het. Stationarity	Robust	Demeaned	-8,30*	0,00
Δnec_t	IPS	Hom. Unit Root	Het. Stationarity	Robust	Demeaned	-34,60*	0,00
nec_t	IPS	Hom. Unit Root	Het. Stationarity	Robust	Demeaned	-7.93*	0,00

Hadri and Im, Pesaran and Shin Panel unit root tests. Following the Dickey and Pantula approach, first differences are tested first. Upon non rejection of the null hypothesis, the test stops, otherwise it continues to levels.

5 Asymmetries - Panel Linearity Tests

Having found evidence of the hysteresis phenomenon, we investigate whether additional nonlinearity could be find in the relationship between such variable and unemployment. Before switching our attention to threshold regressions, we need to test for linearity against a nonlinear alternative through the well established Hansen test as seen in its smoothed variant in Gonzales et al. [16]. We ought to remind that in our estimation strategy, the afore-mentioned test is intended to give us evidence of possible regime-wise nonlinearity, making us feel more comfortable as we proceed to select nonlinear models and comparing them among themselves and the base linear model focusing exclusively on a measure of fit (an information criteria, in our case). In this whole panel pre-test, the test equation, based on an ADF reparametrization of the necessity entrepreneurs series, contains smoothing component which mimics with a logistic function a smooth dummy variable (Equation 7):

$$\Delta nec_{i,t} = \alpha_i + \lambda_t + \beta_1^{'} nec_{i,t-1} + (\beta_2^{''} nec_{i,t-1}) * G(u_{i,t};\gamma;\theta) + \varepsilon_{i,t} \tag{7}$$

the test requires substituting $G(s_{i,t};\gamma;c)$ with its first order Taylor expansion centered around $\gamma = 0$. (as we just explained, Hansen extended to Panel). Generalizing the expansion to up to m regimes, we would get Equation 8:

$$\Delta nec_{i,t} = \mu_i + \beta_0^{'*} nec_{i,t} + \beta_1^{'*} nec_{i,t} q_{i,t} + ... + \beta_m^{'*} nec_{i,t} q_{i,t}^m + u_{it}^* \tag{8}$$

with its null based on the joint nonsignificance of all the coefficients of the branches of the regression, which is equivalent to set the slope of function G equal to 0 in Equation 9

$$H_0 : \beta_1^{'*} = ... = \beta_m^{'*} = 0 \simeq H_0 : \gamma = 0 \tag{9}$$

Given how delicate this first a priori test is in giving us an idea of how much a generic nonlinear alternative could perform against the null of absolute linearity, we choose to bootstrap the test results around one hundred times. Furthermore, up until this point of our work, the test against any homogeneous alternative has been done at the panel level, for an autoregressive (ADF type) specification

Table 11: Linearity tests against one or two regime models. Alternative tests.

nec_t	*m*	*LM(x)*	*p-value*	*LM(F)*	*p-value*	*HAC(X)*	*p-value*	*HAC(F)*	*p-value*
	0vs1	8,303	0,004	8,190	0,004	3,461	0,063	3,414	0,065
	0vs2	12,420	0,002	6,120	0,002	3,470	0,176	1,710	0,181
	0vs1	8,303	0,004	8,190	0,004	3,461	0,063	3,414	0,065
	2vs1	4,139	0,042	4,080	0,044	1,887	0,170	1,860	0,173
opo_t	*m*	*LM(x)*	*p-value*	*LM(F)*	*p-value*	*HAC(X)*	*p-value*	*HAC(F)*	*p-value*
	0vs1	16,250	0,001	16,030	0,001	3,396	0,065	3,350	0,067
	0vs2	16,370	0,001	8,066	0,001	4,862	0,088	2,396	0,091
	0vs1	16,250	0,001	16,030	0,001	3,396	0,065	3,350	0,067
	2vs1	0,117	0,001	0,116	0,001	1,793	0,181	1,767	0,184

Heteroskedasticity consisent LM and F tests.

with an attractor and with the unemployment rate as the switching variable. The battery of tests (LM, F and HAC corrected based on the joint null of each coefficient of each regime (from zero to a maximum of two) generally held favorable indication of possible nonlinearity in the opportunity entrepreneurs group, pointing at the possible existence of at least a pair of regimes. The very same result was detected for the joint panel hypothesis of no statistical difference between segment slopes for the necessity group, with a p value comfortably close to five percent in all five tests we specified, efficiently rejecting the null of absolute linearity at the overall panel level. The results we have just commented on are readily available in Table 11. Even though a time series based Hansen test or any other given test of linearity against general nonlinearity might be used to investigate each and every region at the sectorial level, this outcome is thus granting us enough evidence to proceed with the choice of more specific threshold functional forms at such lower level of detail.

6 Threshold Regressions with unemployment as a threshold

In order to give the reader an idea of the models of choice, let's focus on a generic TAR first, adjusting it to our case: an unrestricted, **three regimes TAR**, but with an ADF specification, would look something like this:

$$\Delta nec_t = \begin{cases} c_1 + \rho_l nec_{t-1} + e_t \text{ if } u_t \leq \theta_l \\ c_2 + \rho_m nec_{t-1} + e_t \text{ if } \theta_l \leq u_t \leq \theta_h \\ c_3 + \rho_h nec_{t-1} + e_t \text{ if } u_t \geq \theta_h \end{cases} \tag{10}$$

With nec_t necessity entrepreneurs at time t, u_t unemployment level, θ threshold value of u, and e_t hopefully (i.i.d.). If only one threshold is detected and considered, such that $\theta_l = \theta_h = \theta$ and $\rho_m = 0$, the model becomes a two branches regression, as showed in Equation 11.

$$\Delta nec_t = \begin{cases} c_1 + \rho_l nec_{t-1} + e_t \text{ if } u_t \leq \theta \\ c_3 + \rho_h nec_{t-1} + e_t \text{ if } u_t \geq \theta \end{cases} \tag{11}$$

Of course, if $\theta_l = \theta_h = 0$, with $\rho_l = \rho_m = \rho_h$ the model collapses to $\Delta nec_t = \rho nec_{t-1} + e_t$, and we are back to the linear case. By any chance, we also would want to specify an unrestricted three regimes SETAR, where the self exciting component is represented by the contemporaneous value of the objective variable. Following on to the example in Equation 10 and targeting the opportunity entrepreneurs series, we would similarly have:

$$\Delta opo_t = \begin{cases} c_1 + \rho_l opo_{t-1} + e_t \text{ if } opo_t \leq \theta_l \\ c_2 + \rho_m opo_{t-1} + e_t \text{ if } \theta_l \leq opo_t \leq \theta_h \\ c_3 + \rho_h opo_{t-1} + e_t \text{ if } opo_t \geq \theta_h \end{cases} \tag{12}$$

6.1 Thresholds autoregressions (Autonomous Communities)

In this section, we run a series of threshold autoregressions on a series of ADF type models of order 1 (obtained by subtracting y_{t-1} to both sides of a random walk with drift process). This way, with (mainly) first differenced stationary variables, the value of the attractor will give us an idea of how fast the recovery is from an exogenous shock (how grave the hysteresis phenomenon is) subject to variations in the unemployment rate, which for now represents our threshold variable. We will evaluate in this case both the level value of the unemployment and its first difference. That will tell us whether the absolute value of unemployment or its short run variation matter more in terms of self employment choice. It is also worth noting that, different from, say, the Hansen test comparing linearity with some non specified form of (regime) nonlinearity through polynomial expansion, in this and the following paragraphs we shall consider goodness of fit as an a posteriori additional rule of choice comparing a linear representation of the decay against a multiple regime alternative. Rather than on the R^2, we shall focus on the BIC as the rule of choice.

Let's first consider unemployment as a threshold: In the case of necessity entrepreneurs, regions Andalucia and Castile and Leon (1 7) presented two distinct threshold (and thus three regimes), regions Asturias, Cantabria and Basque Countries (3 6 16) were considered linear (no threshold could be found) while the rest of the regions showed as a suitable functional form a model with two regimes and a single threshold. Overall, the whole country (18) presents itself with two thresholds and three regimes.

Again, as we consider again unemployment in levels as our threshold variable and we switch our attention to opportunity entrepreneurs, we appear to be able to find a possible three regimes representation for Canary Islands, Murcia, La Rioja and the country as a whole (5 14 17 18), a linear representation for Aragon, Asturias, Cantabria, Valencian Community, Galicia and Navarre (2 3 6 10 12 15) and a two regimes/one threshold representation for the rest of the regions.

Let's now consider the change in unemployment as our objective threshold: as we start with necessity once again, we notice how sensibility of the deviations to changes in unemployment is far lesser compared to sensibility to the overall magnitude of unemployment: only the Balearic Islands (4) would present a model with three regimes and two threshold values, while only Catalonia (9) would show sign of a single threshold value with two regimes. It is however

Table 12: TAR regional models, Necessity entrepreneurs

nec_t	ρ_l	ρ_m	ρ_h	ρ_{linear}
(I)	(II)	(III)	(IV)	(V)
Andalucia	1,047*	-0,399*	-0,889*	-
Aragon	-0,678*	-	-1,018*	-
Asturias	-	-	-	-0,632*
Balearic Islands	-0,778*	-	0,110	-
Canary Islands	-0,424*	-	-1,434*	
Cantabria	-	-	-	0,936*
Castile and Leon	-1,198*	-0,821*	-1,097*	-
Castile-La Mancha	-0,464*	-	-0,878*	-
Catalonia	-0,256*	-	-0,929*	-
Valencian Community	-0,645*	-	-1,107*	-
Extremadura	-0,438*	-	-1,191*	-
Galicia	-1,299*	-	-0,213*	-
Community of Madrid	-0,340*	-	-1,032*	-
Region of Murcia	-0,875*	-	-1,268*	-
Navarre	-0,735*	-	-1,370*	-
Basque Country	-	-	-	-0,795*
La Rioja	-0,966*	-	-1,007*	-
SPAIN	-0,147*	-1,206*	-0,882*	-

Estimates for the best BIC performing two, three regime TAR or linear model. * indicates significant estimates at 5 percent.

Table 13: TAR regional models, Opportunity entrepreneurs

opo_t	ρ_l	ρ_m	ρ_h	ρ_{linear}
(I)	(II)	(III)	(IV)	(V)
Andalucia	-1,152*	-	-0,495*	-
Aragon	-	-	-	-0,437*
Asturias	-	-	-	-0,517*
Balearic Islands	-0,450*	-	-1,179*	-
Canary Islands	-0,297*	-1,411*	-0,522*	-
Cantabria	-	-	-	0,723*
Castile and Leon	-1,285*	-	-0,408*	-
Castile-La Mancha	-0,130	-	-0,792*	-
Catalonia	-0,223*	-	-0,879*	-
Valencian Community	-	-	-	-0,365*
Extremadura	-0,465*	-	-0,798*	-
Galicia	-	-	-	-0,412*
Community of Madrid	-0,281*	-	-1,106*	-
Region of Murcia	-0,403*	-0,528*	-1,287*	-
Navarre	-	-	-	-0,965*
Basque Country	-1,168*	-	-0,549*	-
La Rioja	-0,475*	0,610*	-1,106*	-
SPAIN	-0,550*	-0,319	-0,768*	-

Estimates for the best BIC performing two, three regime TAR or linear model. * indicates significant estimates at 5 percent.

worth noting that the country as a whole (18) would present signs of nonlinearity, although bordering non-significance in the lower regime.

Finally, in the case of the opportunity entrepreneurs, Andalucia, Murcia and Basque Country (1 14 16) would present a three regimes model, Aragon, Asturias, Balearic Islands, Canary Islands, Cantabria, Castile and Leon, Extremadura, Navarre, La Rioja and Spain as a whole (2 3 4 5 6 7 11 15 17 18) would be better represented by a linear model, while the rest would indicate a two regimes model as the best suitable fit for the data.

6.2 Threshold autoregressions (Sectorial Level)

Before discussing the deviation values, it would be useful getting an equivalent overview at the sectorial level. As we base our results on the previous section, we focus on the unemployment rate as a level threshold variable. Starting with the necessity entrepreneurs group, we found a suitable two regime representation for Andalucia, Balearic Islands, Cantabria, Catalonia, Valencian Community, Murcia, Basque Country and Spain as a whole (1 4 6 9 10 14 16

Table 14: Necessity workers, agricultural sector, all fitted models

	ρ_l	ρ_m	ρ_h	ρ_{linear}
(I)	(II)	(III)	(IV)	(V)
Andalucia	-1,251*	-	-1,117*	-
Aragon	-	-	-	-0,880*
Asturias	-	-	-	-0,883*
Balearic Islands	-1,061*	-	-0,897*	-
Canary Islands	-	-	-	-1,082*
Cantabria	-0,685*	-	-1,088*	-
Castile and Leon	-0,512*	-1,378*	-1,091*	-
Castile-La Mancha	-0,482*	-2,699*	-0,967*	-
Catalonia	-1,286*	-	-1,014*	-
Valencian Community	-0,866*	-	-0,861*	-
Extremadura	-	-	-	-0,583*
Galicia	-1,100*	-0,213*	-0,569*	-
Community of Madrid	-	-	-	-0,812*
Region of Murcia	-1,291*	-	-1,010*	-
Navarre	-0,701*	-1,52*	-1,039*	-
Basque Country	-0,583*	-	-0,908*	-
La Rioja	-	-	-	-0,844*
SPAIN	-0,430*	-	-0,455*	-

Estimates for the best BIC performing two, three regime TAR or linear model. * indicates significant estimates at 5 percent.

18) and a suitable three regime representation for Castile and Leon, Castile and La Mancha, Galicia and Navarre (7 8 12 15) in the agricultural sector. As we moved on to the industry sector, we found a good fitness for a two regime model in Castile and La Mancha and Catalonia (8 9) and a three regime model in Extremadura, Community of Madrid, Basque Country and Spain as a whole (11 13 16 18). Moving on to the construction sector, regions Aragon, Balearic Islands, Castile and Leon, Castile and La Mancha, Valencian Community, Galicia, Murcia, Navarre, La Rioja and Spain as a whole (2 4 7 8 10 12 14 15 17 18) were found to fit better two regimes, while no regions could be more efficiently described by a 3 regimes model. In the service sector, the low qualified group presented a satisfying fitness, although restricted to two regimes only, for Balearic Islands, Canary Islands, Castile and Leon, Catalonia, Galicia, Community of Madrid, Murcia (4 5 7 9 12 13 14), while the three regimes one would be a valid alternative for Andalucia, Castile and La Mancha, Valencian Community, Extremadura, Basque Country and Spain as a whole (1 8 10 11 16 18). Finally, the high qualified service group had regions Andalucia, Aragon, Canary Islands, Cantabria, Castile and La Mancha, Catalonia, Murcia and Navarre (1 2 5 6 8 9 14 15) well described by a two regimes transition model, and Balearic Islands, Valencian Community, Community of Madrid and the whole country (4 10 13 18) by a three regimes model. It would thus appear the service sector would present much of the asymmetric variation and adjustment we would see in the bust and boom cycle of necessity self-employment.

Let's now switch our attention to the opportunity entrepreneurs group. As for the agricultural sector, Andalucia, Cantabria, Castile and La Mancha, Catalonia, Valencian Community, Extremadura, Galicia, Community of Madrid, Murcia, Basque Country, La Rioja and Spain as a whole (1 6 8 9 10 11 12 13 14 16 17 18) would show a suitable two regime modelling solution, while Asturias (3) only would point at a three regimes structure. In the industrial sector, regions Andalucia, Aragon, Castile and Leon, Castile and La Mancha, Catalonia, Extremadura, Galicia, Murcia, Navarre and Basque Country (1 2 7 8 9 11 12 14 15 16) would suggest a two regimes structure, while Valencian Community and

Table 15: Necessity workers, Industry sector, all fitted models

	ρ_l	ρ_m	ρ_h	ρ_{linear}
(I)	(II)	(III)	(IV)	(V)
Andalucia	-	-	-	-1,097*
Aragon	-	-	-	-1,084*
Asturias	-	-	-	-1,025*
Balearic Islands	-	-	-	-1,084*
Canary Islands	-	-	-	-0,957*
Cantabria	-	-	-	-0,734*
Castile and Leon	-	-	-	-0,816*
Castile-La Mancha	-0,323*	-	-1,065*	-
Catalonia	-1,287*	-	-0,897*	-
Valencian Community	-	-	-	-1,086*
Extremadura	-0,731*	-0,291	-1,066*	-
Galicia	-	-	-	-0,764*
Community of Madrid	-0,659*	-1,768*	-1,263*	-
Region of Murcia	-	-	-	-1,094*
Navarre	-	-	-	-1,075*
Basque Country	-1,025*	-1,150*	-1,352*	-
La Rioja	-	-	-	-1,032*
SPAIN	-0,735*	-0,508*	-1,541*	-

Estimates for the best BIC performing two, three regime TAR or linear model. * indicates significant estimates at 5 percent.

Table 16: Necessity workers, Construction sector, all fitted models

	ρ_l	ρ_m	ρ_h	ρ_{linear}
(I)	(II)	(III)	(IV)	(V)
Andalucia	-	-	-	-0,598
Aragon	2,622*	-	-1,041*	-
Asturias	-	-	-	-0,852*
Balearic Islands	-1,033*	-	-1,285*	-
Canary Islands	-	-	-	-0,842*
Cantabria	-	-	-	-0,926*
Castile and Leon	-1,076*	-	-0,992*	-
Castile-La Mancha	-0,917*	-	-0,861*	-
Catalonia	-	-	-	-0,581*
Valencian Community	-0,852*	-	-1,355*	-
Extremadura	-	-	-	-1,080*
Galicia	-1,108*	-	-0,634*	-
Community of Madrid	-	-	-	-0,869*
Region of Murcia	-0,971*	-	-0,948*	-
Navarre	-1,132*	-	-1,117*	-
Basque Country	-	-	-	-1,145*
La Rioja	-0,940*	-	-1,048*	-
SPAIN	-0,364*	-	-0,847*	-

Estimates for the best BIC performing two, three regime TAR or linear model. * indicates significant estimates at 5 percent.

Table 17: Necessity workers, Low skilled Service sector, all fitted models

	ρ_l	ρ_m	ρ_h	ρ_{linear}
(I)	(II)	(III)	(IV)	(V)
Andalucia	-1,379*	-0,668*	-0,674*	-
Aragon	-	-	-	-0,908*
Asturias	-	-	-	-0,975*
Balearic Islands	-0,832*	-	-0,463	-
Canary Islands	-0,871*	-	-1,237*	-
Cantabria	-	-	-	-1,145*
Castile and Leon	-1,204*	-	-0,973*	-
Castile-La Mancha	-0,674*	-1,421*	-0,841*	-
Catalonia	-0,298*	-	0,739*	-
Valencian Community	-1,065*	-1,394*	-0,812*	-
Extremadura	-0,944*	-1,582*	-1,586*	-
Galicia	-1,384*	-	-0,679*	-
Community of Madrid	-0,742*	-	-0,914*	-
Region of Murcia	-0,993*	-	-1,337*	-
Navarre	-	-	-	-1,058*
Basque Country	-0,947*	-	-1,358*	-
La Rioja	-	-	-	-0,985*
SPAIN	-0,314*	-1,499*	-0,780*	-

Estimates for the best BIC performing two, three regime TAR or linear model. * indicates significant estimates at 5 percent.

Table 18: Necessity workers, High skilled Service sector, all fitted models

	ρ_l	ρ_m	ρ_h	ρ_{linear}
(I)	(II)	(III)	(IV)	(V)
Andalucia	-0,445*	-	-1,170*	-
Aragon	-0,908*	-	0,999*	-
Asturias	-	-	-	-0,992*
Balearic Islands	-0,799*	-0,937*	-0,548*	-
Canary Islands	-0,700*	-	-1,521*	-
Cantabria	-0,853*	-	-1,827*	-
Castile and Leon	-	-	-	-1,028*
Castile-La Mancha	-1,103*	-	-0,619*	-
Catalonia	-0,624*	-	-1,467*	-
Valencian Community	-0,504*	-0,842*	-0,994*	-
Extremadura	-	-	-	-0,961*
Galicia	-	-	-	-0,770*
Community of Madrid	-0,319*	-1,138*	-1,158*	-
Region of Murcia	-1,184*	-	-0,975*	-
Navarre	-0,869*	-	-1,949*	-
Basque Country	-	-	-	-0,978*
La Rioja	-	-	-	-1,115*
SPAIN	-0,693*	-0,269	-1,271*	-

Estimates for the best BIC performing two, three regime TAR or linear model. * indicates significant estimates at 5 percent.

Spain as a whole (10 18) would better welcome a three regimes model. In the constructions sector, regions Andalucia, Aragon, Castile and Leon, Castile and La Mancha, Catalonia, Valencian Community, Extremadura, Galicia, Murcia, Navarre, Basque Country and Spain as a whole (1 2 7 8 9 10 11 12 14 15 16 18) would fit better two regimes model, while Canary Islands (5) only would better fit a model with two threshold values and two regimes. Switching to the low skilled service sector, regions Castile and Leon, Castile and la Mancha, Community of Madrid, Navarre, Basque Country and La Rioja (7 8 13 15 16 17) would present a two regimes model, while regions Andalucia, Balearic Islands, Cantabria, Catalonia, Valencian Community, Murcia and Spain as a whole (1 4 6 9 10 14 18) would find a three regimes ones a better solution. This last one is perhaps the sector where the number of transitions equals 2 in more regions than those where it equals only one: that would point out at the existence of intermediate values of unemployment for which the rate of return of the unconditional mean of self-employment either speeds down or is not even present (quasi unit root behavior in the internal regime). Finally, as we have a look at the service sector share compose of high skilled labor, we observe that Balearic Islands, Castile and La Mancha, Community of Madrid and Murcia (4 8 13 14) would accept a two regimes representation, while regions Andalucia, Canary Islands, Catalonia, Valencian Community, La Rioja and Spain as a whole (1 5 9 10 17 18) would go for a three regimes one. It is perhaps important to point out how the sector with the smoother transitions (mind, in terms of numbers of regimes/threshold values, and not how sharp the transition is), is perhaps the service sector. Economic actors in the service market tend to adapt less abruptly to market news than their agricultural and industrial equivalents.

Table 19: Opportunity workers, agricultural sector, all fitted models

	ρ_l	ρ_m	ρ_h	ρ_{linear}
(I)	(II)	(III)	(IV)	(V)
Andalucia	-1,053*	-	-0,730*	-
Aragon	-	-	-	-0,867*
Asturias	-0,965*	0,352	-1,043*	-
Balearic Islands	-	-	-	-1,043*
Canary Islands	-	-	-	-0,985*
Cantabria	-0,480*	-	-1,199*	-
Castile and Leon	-	-	-	-0,692*
Castile-La Mancha	-0,575*	-	-1,081*	-
Catalonia	-0,753*	-	-0,750*	-
Valencian Community	-0,545*	-	-1,178*	-
Extremadura	-0,592*	-	-0,740*	-
Galicia	-0,377*	-	-1,088*	-
Community of Madrid	-0,966*	-	-1,007*	-
Region of Murcia	-0,766*	-	-1,366*	-
Navarre	-	-	-	-0,940*
Basque Country	-1,423*	-	-0,974*	-
La Rioja	-0,870*	-	-1,537*	-
SPAIN	-0,363*	-	-0,803*	-

Estimates for the best BIC performing two, three regime TAR or linear model. * indicates significant estimates at 5 percent.

Table 20: Opportunity workers, industry sector, all fitted models

	ρ_l	ρ_m	ρ_h	ρ_{linear}
(I)	(II)	(III)	(IV)	(V)
Andalucia	-1,421*	-	-1,066*	-
Aragon	-0,727*	-	-1,013*	-
Asturias	-	-	-	-0,809*
Balearic Islands	-	-	-	-1,037*
Canary Islands	-	-	-	-1,015*
Cantabria	-	-	-	-1,038*
Castile and Leon	-1,824*	-	-0,676*	-
Castile-La Mancha	-0,210	-	-0,885*	-
Catalonia	-0,794*	-	-1,056*	-
Valencian Community	-1,818*	0,451	-1,096*	-
Extremadura	-0,930	-	-0,986*	-
Galicia	-0,658*	-	-1,014*	-
Community of Madrid	-	-	-	-1,045*
Region of Murcia	-0,700*	-	-1,178*	-
Navarre	-1,018*	-	-0,755*	-
Basque Country	-1,202*	-	-0,756*	-
La Rioja	-	-	-	-0,699*
SPAIN	-0,822*	0,294	-0,941*	-

Estimates for the best BIC performing two, three regime TAR or linear model. * indicates significant estimates at 5 percent.

Table 21: Opportunity workers, construction sector, all fitted models

	ρ_l	ρ_m	ρ_h	ρ_{linear}
(I)	(II)	(III)	(IV)	(V)
Andalucia	-1,294*	-	-0,754*	-
Aragon	-0,588*	-	-1,006*	-
Asturias	-	-	-	-0,721*
Balearic Islands	-	-	-	-0,985*
Canary Islands	-0,700*	-1,289*	-1,147*	-
Cantabria	-	-	-	-0,856*
Castile and Leon	'-0,851*	-	-0,674*	-
Castile-La Mancha	-0,659*	-	-0,845*	-
Catalonia	-0,685*	-	-0,909*	-
Valencian Community	-0,333*	-	-1,015*	-
Extremadura	-1,583*	-	-0,833*	-
Galicia	-0,945*	-	-0,813*	-
Community of Madrid	-	-	-0,882*	-
Region of Murcia	-0,988*	-	-0,679*	-
Navarre	-1,408*	-	-0,901*	-
Basque Country	-0,885*	-	-0,939*	-
La Rioja	-	-	-	-0,978*
SPAIN	-0,614*	-	-0,400*	-

Estimates for the best BIC performing two, three regime TAR or linear model. * indicates significant estimates at 5 percent.

Table 22: Opportunity workers, Low skill Service sector, all fitted models

	ρ_l	ρ_m	ρ_h	ρ_{linear}
(I)	(II)	(III)	(IV)	(V)
Andalucia	-0,797*	-0,726*	-1,120*	-
Aragon	-	-	-	-0,649*
Asturias	-	-	-	-0,579*
Balearic Islands	-0,812*	-0,501	-1,224*	-
Canary Islands	-	-	-	-0,808*
Cantabria	-0,835*	-0,481*	-0,915*	-
Castile and Leon	-0,291*	-	-0,947*	-
Castile-La Mancha	-0,253*	-	-0,898*	-
Catalonia	-00262*	-1,331*	-0,983*	-
Valencian Community	-0,448	-0,945*	-0,769*	-
Extremadura	-	-	-	-0,774*
Galicia	-	-	-	-0,593*
Community of Madrid	-0,531*	-	-1,275*	-
Region of Murcia	-0,622*	-0,975*	-1,349*	-
Navarre	-1,476*	-	-1,025*	-
Basque Country	-1,333*	-	-0,669*	-
La Rioja	-0,437*	-	-1,104*	-
SPAIN	-0,582*	-0,367	-1,135*	-

Estimates for the best BIC performing two, three regime TAR or linear model. * indicates significant estimates at 5 percent.

Table 23: Opportunity workers, High skill Service sector, all fitted models

	ρ_l	ρ_m	ρ_h	ρ_{linear}
(I)	(II)	(III)	(IV)	(V)
Andalucia	-1,127*	-1,371	-1,025*	-
Aragon	-	-	-	-0,746*
Asturias	-	-	-	-0,802*
Balearic Islands	-0,418*	-	-1,106*	-
Canary Islands	-0,461*	-1,591*	-0,842*	-
Cantabria	-	-	-	-0,869*
Castile and Leon	-	-	-	-0,736*
Castile-La Mancha	-0,416*	-	-1,353*	-
Catalonia	-0,251*	-1,072*	-1,027*	-
Valencian Community	-0,589*	-1,205*	-0,946*	-
Extremadura	-	-	-0,970*	-
Galicia	-	-	-	-0,656*
Community of Madrid	-0,554*	-	-1,110*	-
Region of Murcia	-0,348*	-	-0,951*	-
Navarre	-	-	-	-0,784*
Basque Country	-	-	-	-0,715*
La Rioja	-0,707*	0,276	-1,300*	-
SPAIN	-0,582*	-0,162	-1,060*	-

Estimates for the best BIC performing two, three regime TAR or linear model. * indicates significant estimates at 5 percent.

7 Inference on the half life of shocks to employment

7.1 Rate of decay of the shocks to self-employment, national level

We now calculate the rate of decay (λ, which amounts to ρ expressed in terms of time) of an autoregressive process of order 1 with a lagged attractor from the series at the regional level. Considering our previous results, it appears more than sensible to offer the calculated time deviations based on the level threshold of unemployment.

Given the results obtained on Spain (18) at the most aggregated possible level in the previous section, we will discuss and calculate the rate of decay of the process for necessity and opportunity entrepreneurs with a three regimes structures for both, and compare it to the linear specification in each case. We remind the reader that, given the simple ADF specification of our test equation, the inverse formula required to calculate the half lives of the deviations will be: $\lambda^T = (1 - x)/(1 - \lambda)$.

Beginning with the necessity entrepreneurs, our significant coefficients of error corrections for the three branches (regimes) are thus: -0.15, -1.21 and -0.88 in the low, middle and high regime of unemployment. This corresponds, in the two most extreme regimes, to a decay of 1.70 and 0.24 quarters respectively, required to recover half the positive shock to unemployment the economy has sustained. Evidently, switching from a lower state to a higher state of unemployment causes three effects: as unemployment is lower than its "natural rate", after a positive shock which increases the number of necessity entrepreneurs, the rate of adjustment back to equilibrium is relatively slow. As unemployment increases and gets close to its natural rate (say, the middle regime), the speed of adjustment starts to accelerate as deviations are compensated more quickly (the economy and its operators have had time to learn how to react to the "natural rate of unemployment"). However, as the unemployment rate increases above its calculated threshold value (which we have improperly called up until now "natural rate") the speed of adjustment after a shock starts to slow down again, as more and more self employed operators cannot be reabsorbed once more into the economy as salary workers. The richness of the description offered by the model clashed with the linear value of the attractor, whose coefficient is close to 0,22.

On a similar note, but with a central region behaving with a bit of a unit root behavior the values for the significant lower and upper regime of the opportunity entrepreneurs model are -0.55, and -0.77 respectively. The speeds of adjustment, which in quarters amount to a half life of 0,90 and 0,46, follow a very similar story when compared to the necessity entrepreneurs group, with a higher half life in the high unemployment regime with respect to the low unemployment regime: this suggests us that pull factors might be far stronger than push factors in prolonged states of above the level unemployment. Pretty much in line with

Table 24: All sectors, national level, all fitted models

Variable	*Regimes*			*Linear*
nec_t	ρ_l	ρ_m	ρ_h	ρ_{linear}
Agriculture	-0,430*	-	-0,455*	-
Industry	-0,735*	-0,508*	-1,541*	-
Construction	-0,364*	-	-0,847*	-
Low Service	-0,314*	-1,499*	-0,780*	-
High Service	-0,693*	-0,269	-1,271*	-
opo_t	ρ_l	ρ_m	ρ_h	ρ_{linear}
Agriculture	-0,363*	-	-0,803*	-
Industry	-0,822*	0,294	-0,941*	-
Construction	-0,614*	-	-0,400*	-
Low Service	-0,582*	-0,367	-1,135*	-
High Service	-0,582*	-0,162	-1,060*	-

National level regime analysis. Estimates for the best BIC performing two, three regime TAR or linear model. * indicates significant estimates at 5 percent.

the linear ADF model for the necessity entrepreneurs, the speed of adjustment in the case of the opportunity entrepreneurs is again fixed around 0,22. That amounts to a rate of decay of around 1,54 quarters.

7.2 Rate of decay of the shocks to self-employment, sectorial level

How does the situation differ across sectors? A relevant achievement of the convergence analysis at the sectorial level, which is visible in Table 24, is that in no case a linear alternative was deemed superior to a regime model in any specification across all sectors when goodness of fit was compared via the Bayesian Information criteria: the analysis shows how across all sectors two to three regimes where considered and always performed better than the linear alternative. Focusing on the opportunity entrepreneurs group, Agriculture and Construction where the only to sectors to be better suited by a two regime model, while Industry and both service sectors pointed at three regime modelling. In particular, and similarly to the result we obtained at the national level in Table 25, the former sectors showed the presence of what the literature on persistence would consider a form of "inaction band": as the coefficient governing the middle regime[2] is statistically imprecise and thus not so distant from 0, the opportunity entrepreneurs across sectors would not converge neither diverge to its mean as the model would collapse, for $\rho_m = 0$, to the benchmark random walk with drift. This would mean that, for values of unemployment inside such a "natural interval " in the middle of the regression model, the model would act in a quasi-unit root manner, and wander aimlessly as time progresses. The rest of the inference base on the values of the branches which appear statistically significant do not appear to tell us a story any different from the aggregated case: higher speed of adjustment back to the equilibrium in the upper regime, where unemployment is at a higher threshold level, slower speed of adjustment in the lower regime, the only possible exception made for constructions where the opposite would happen (-0,614* in the lower regime as opposed to -0,400* in the upper regime).

[2]-0,31(0,18) at the national level, close to such value at the Sectorial level, ranging from -0,162 in High Skill Services to -0,367 in Low Skill Services.

Table 25: Speed of convergence estimates, AR(1) and BEST TAR

	$ADF(1)$	TAR		
	ρ	$\rho_{l,quarters}$	$\rho_{m,quarters}$	$\rho_{h,quarters}$
(I)	(II)	(III)	(IV)	(V)
Δnec_t		-0,15*	-1,21*	-0,88*
		(0,07)	(0,23)	(0,27)
Δnec_t	-0,22			
	(0,07)			
Δopo_t		-0,55*	-0,31	-0,77*
		(0,10)	(0,18)	(0,27)
Δopo_t				
	-0,23			
	(0,07)			

Estimates for the best BIC performing two, three regime TAR or linear model. * indicates significant estimates at 5 percent.

Table 26: Half life estimates, AR(1) and BEST TAR

	$ADF(1)$	TAR		
	λ	$\lambda_{l,quarters}$	$\lambda_{m,quarters}$	$\lambda_{h,quarters}$
(I)	(II)	(III)	(IV)	(V)
Δnec_t		1,70	0.42	0,24
Δnec_t	1,56			
Δopo_t		0,90	-	0,46
Δopo_t	1,54			

Half-life estimates of the attractors/error correction parameters, calculated according to $\lambda^T = (1-x)/(1-\lambda)$.

8 Threshold regressions with GDP as a threshold

The analysis showed until here strictly focused on the Unemployment rate as an exogenous threshold. However, to understand and countercheck the feasibility of a nonlinear representation of entrepreneur variations across the necessity and opportunity group, we also tested for a GDP proxy variable to corroborate our results. As opportunity entrepreneurs in particular take note of boom and busts of the business cycle and try to take advantage of news about the expansion phase of the cycle as a mean to fire up their business, they will more than likely monitor a variable such as an industrial production index. In this section we report and comment results on TAR model specifications when the threshold variable in an index of the kind mentioned above. Although the same reasoning does not exactly apply to necessity entrepreneurs, who would become such in case of economic need, thus in case of bad news about employment or during low phases of the cycles, we could perhaps argue, consistently with literature, that opportunity workers would follow prociclically the economy, while necessity workers would present an anticiclycal behavior. Whatever the case, we should thus expect necessity entrepreneurs numbers to decrease or increase more sluggishly during boom phases, as they would be content with their wage and be effectively employed, thus reacting slowly to any sort of exogenous shock pushing them away from their natural average number. On the opposite side, deviations from their average number in the case of opportunity workers would imply, in the case of an external shock and everything else equal, a faster speed of adjustment to their mean, even more so when the shock is negative and opportunity entrepreneurs numbers have been reduced by it. Lastly, we present

Table 27: Opportunity workers, GDP as threshold, all fitted models

	ρ_l	ρ_m	ρ_h	ρ_{linear}
(I)	(II)	(III)	(IV)	(V)
Andalucia	-0,831*	-1,509*	-0,327*	
Aragon	-1,035*		-0,626*	
Asturias	-0,575*		-0,776*	
Balearic Islands	-1,356*		-0,950*	
Canary Islands	-0,708*		-0,761*	
Cantabria	-0,882*		-0,817*	
Castile and Leon				-0,342*
Castile-La Mancha	-0,767*		-0,545*	
Catalonia	-0,437*		-0,874*	
Valencian Community	-0,315*		-0,676*	
Extremadura				-0,462*
Galicia	-0,942*	-0,640*	-0,725*	
Community of Madrid	-0,692*	-1,360*	-1,110*	
Region of Murcia				-0,592*
Navarre	-1,091*	-1,380*	-1,103*	
Basque Country	-0,821*	-0,736*	-0,873*	
La Rioja	-0,376*		-0,986*	
SPAIN	-0,569*	-0,760*	-0,299*	

Estimates for the best BIC performing two, three regime TAR or linear model. * indicates significant estimates at 5 percent.

Table 28: Necessity workers, GDP as threshold, all fitted models

	ρ_l	ρ_m	ρ_h	ρ_{linear}
(I)	(II)	(III)	(IV)	(V)
Andalucia	-0,082*	-0,789*	-1,148*	
Aragon	-0,936*	-0,976*	-1,053*	
Asturias	-0,352*		-0,857*	
Balearic Islands	0,105		-1,017*	
Canary Islands	0,619		-0,855*	
Cantabria				-0,936*
Castile and Leon				-0,711*
Castile-La Mancha	-0,956*	-0,900*	-0,631*	
Catalonia	0,390		-0,891*	
Valencian Community	-0,220*	-1,289*	-0,960*	
Extremadura	-0,967*	-1,064*	-0,977*	
Galicia	-1,036*	-0,400*	-0,942*	
Community of Madrid	-0,695*	-1,118*	-0,883*	
Region of Murcia	-1,272*	-1,166*	-1,246*	
Navarre	-1,104*	-1,094*	-0,963*	
Basque Country	0,442*		-1,274*	
La Rioja	-1,010*		-1,006*	
SPAIN	-0,928*	-0,964*	-0,801*	

Estimates for the best BIC performing two, three regime TAR or linear model. * indicates significant estimates at 5 percent.

estimates on the opportunity workers speed of adjustment with a one period lagged GDP variable. This represents an important refinement in the procedure, crucially important at higher frequencies, such as monthly or weekly ones. Even though we have been working with quarterly observations, we might assume that a spurious relationship might exist between our threshold variables and the entrepreneurs groups in the contemporaneous model, as economic actors need some time to adjust to information and react to it. Lagging the threshold variable could perhaps lead to either a confirmation of some degree of nonlinear adjustment or push our conclusions back in favor of a linear rendition of the model.

Results for the threshold regression with the industrial production index as a threshold are visible in Tables 28 and 27. For the necessity entrepreneurs, the adoption of a GDP measure halves the number of regions where no hint of regime behavior was found, as only Extremadura, Murcia and Castile and Leon (7 11 14) where found to not reject the feasibility of a linear representation. As for the opportunity entrepreneurs group as a whole, the industrial index also managed to reduce the number of linear representations by one, leaving

Table 29: Opportunity workers, lagged GDP as threshold, all fitted models

	ρ_l	ρ_m	ρ_h	ρ_{linear}
(I)	(II)	(III)	(IV)	(V)
Andalucia	-0,945*	-2,693*	-0,328*	
Aragon	-0,675*		-0,627*	
Asturias	-0,651*		-0,794*	
Balearic Islands	-1,021*		-0,846*	
Canary Islands	-0,794*		-0,798*	
Cantabria	-0,929*	-0,946*	-0,885*	
Castile and Leon	-0,885*		-0,497*	
Castile-La Mancha	-0,802*	-0,713*	-0,519*	
Catalonia	-0,142*	-0,660*	-1,364*	
Valencian Community	-0,330*		-0,612*	
Extremadura	-0,855*	-1,045*	-0,645*	
Galicia	-0,941*	-0,739*	-0,684*	
Community of Madrid	-0,692*	-1,089*	-1,166*	
Region of Murcia	-0,499*		-1,335*	
Navarre	-1,191*	-1,519*	-1,145*	
Basque Country	-1,077*	-0,786*	-0,879*	
La Rioja	0,431*		-0,986*	
SPAIN	-0,366*	-0,792*	-0,311*	

Estimates for the best BIC performing two, three regime TAR or linear model. * indicates significant estimates at 5 percent.

only Navarre and Castile and Leon inside the linear representation. Overall results at the national level show slightly different pictures when the two distinct threshold are compared: in the case of the opportunity group, the selected two regime model becomes a three regimes one, with the middle regime, indicating an average (natural) GDP index, pointing at the fastest adjustment (-0,760*) where in the case of the unemployment threshold, a similar value was found to be in the high unemployment regime (-0,768*). This is probably due to the cautionary behavior of entrepreneurs when GDP exceeds its trending behavior and will be checked for in the next subsection, when we will present the results for the robustness test with the lagged GDP threshold. As for the necessity entrepreneurs at the national level, the accepted model is still a three regimes one, but perhaps consistently with a situation where high unemployment mirrors low GDP, the convergence rate of the numerosity of the group after an exogenous shock, everything else equal, is perhaps much faster in the low regime with GDP as a threshold (-0,928* compared to -0,147*) and sits more closely to the adjustment value with high levels of unemployment (-0,882*).

8.1 Robustness check: lagged GDP as a threshold variable for the opportunity entrepreneurs group

In the previous section threshold regression with an industrial index proxying GDP was found to greatly increase the performance and eligibility of nonlinear modelling at the regional level, halving the number of linear models left to just three regions. As we have explained at the beginning of the main section, it is very likely that entrepreneurs, especially those seeking profits and seeing real value in their enterprise, would plan ahead their moves and strategies and make forecasts on a set of economic indicators of their choosing in any given moment new data is ultimately available, before resetting their expectations given the arrival of additional information. In this sense, we decided to lag one quarter behind our GDP variable and see how such threshold would behave in affecting the adjustment of the inflows and outflows of the group compared to the

Table 30: Half life estimates, Opportunity Entrepreneurs, GDP vs lagged GDP

	TAR		
	$\lambda_{l,quarters}$	$\lambda_{m,quarters}$	$\lambda_{h,quarters}$
(I)	(II)	(III)	(IV)
$\Delta opo_t, GDP_{t-1}$	0,366	0,279	0,381
$\Delta opo_t, GDP_t$	0,319	0,284	0,385

Half-life estimates of the attractors/error correction parameters, calculated according to $\lambda^T = (1-x)/(1-\lambda)$.

previous contemporaneous applications. Results for this simple robustness tests are shown in Table 29. As our choice of GDP as a threshold given the arguably procyclical nature of the opportunity workers group allowed us to halve the number of regions pointing at linear modelling as the best fitted choice, choosing to lag the indicator one period allowed us to obtained the same boost in terms of nonlinear representation of the regional convergence rates. Strikingly, nonlinear threshold models are thus accepted in every single Spanish region. In Table 30 we finally offer a direct comparison between the half lives of deviations in the contemporaneous and lag models for the opportunity entrepreneurs group. The rate of decay of any given exogenous shock goes between a minimum of 25 to a maximum of 35 days in each state, the time required to halve the impact of the shock is as such relatively low regardless of the current level of GDP but relatively faster when GDP is in its low state (0,366 compared to 0,381, that is, 33 days compared to 34 in the GDP lagged model and 0,319 against 0,385 in the contemporaneous model, which would amount to 29 days against 35). Economic actors react to external shocks in the low state with a higher degree of urgency.

9 Transition probabilities and Markov regressions

9.1 Brief notes on Markov Switching

Up until this very moment, our exercise has focused on checking for the existence of the hysteresis phenomenon and testing whether or not some form of regime driven nonlinearity might be accepted in our entrepreneurial framework. As a last addition to our work, we now ask what is the likelihood, given an a priori information set, for the numerosity of the opportunity and necessity entrepreneurs group to persist around a given level before switching to another value. Such a question calls for a simple application of Markov switching behavior in our time series context.

We remind the reader that a Markov Switching model is a widely use regime switching regression model used in many applications, such as in the seminal paper by Hamilton (1989) [17] where the author used it to capture the different dynamics of the growth rate of GDP during the boom and bust phases of the economic cycle, extending Markov Switching to autoregressive models. Besides its use to untangle the asymmetric evolution of GDP across expansions and recessions, the methodology has been applied in other contexts, from international economics to health sciences and psychology. In our exercise, we apply an autoregressive Markov switching model (ARMS) to the Spanish national level series for opportunity and necessity entrepreneurs. The choice of an autoregressive set-up, besides its convenience, is justified by the fact that it suits gradual adjustments better than the standard Markov switching dynamic regression model (DRMS) and is perhaps more useful to our study since we employ it to check for regime changes in the numerosity of our variables regardless of any weakly exogenous component that might govern such nonlinearity. This last point is enhanced by the nature of the model itself, where the different states of the world follow a Markov process and are naturally unobserved.

$$\begin{aligned} y_t \quad = \quad & \mu_{s_t} + x_t\alpha + z_t\beta_{s_t} \\ & +\phi_{1,s_t}(y_{t-1} - \mu_{s_{t-1}} - x_{t-1}\alpha - z_{t-1}\beta_{s_{t-1}}) \\ & +\phi_{2,s_t}(y_{t-2} - \mu_{s_{t-2}} - x_{t-2}\alpha - z_{t-2}\beta_{s_{t-2}}) \\ & +\varepsilon_{s_t} \end{aligned} \tag{13}$$

In its classic rendition, following Hamilton, an ARMS(2) model closely resembles Equation 13: in this model y_t is the time dependent variable; μ_{s_t} the state/time dependent uncoditional mean; x_t is a set of time dependent covariates whose coefficients do not depend on time and more importantly on space; z_t is a set of covariates whose coefficients are state dependent; ϕ_{1,s_t} ϕ_{2,s_t} are the two autoregressive coefficients in state s_t. In our application, we have two states (s=2, mirroring the threshold two regimes regression) and no additional covariates x_t or z_t are present. As for the variance and the autocorrelation of the process, we tested for both a constant variance ($\sigma^2_{s=1} = \sigma^2_{s=2} = \sigma^2$) and a changing one ($\sigma^2_{s=1} \neq \sigma^2_{s=2}$) across states. The same for the auotregressive components, which were fixed and then let free to shift across states. A simplified baseline model, with again two autoregressive parameters, is a such described in Equation 14.

$$y_t = \mu_{s_t} + \phi_{1,s_t}(y_{t-1} - \mu_{s_{t-1}}) + \phi_{2,s_t}(y_{t-2} - \mu_{s_{t-2}}) + \varepsilon_{s_t} \tag{14}$$

In our application, we need to point out, up to two autoregressive parameters were tested: as a rule of thumb, given we are facing quarterly data and to keep the analysis tractable, considering a value close to Schwert maximum length rule for maximum lags ($[12(\frac{T}{100})]^{1/4} \simeq 1,8$), we opted to test for no more than that. Obviously, if only one lag is selected, the mode collapses to Equation 15.

$$y_t = \mu_{s_t} + \phi_{1,s_t}(y_{t-1} - \mu_{s_{t-1}}) + \varepsilon_{s_t} \tag{15}$$

9.2 Application and Results

We now report the results of the Markov estimates on the national Spanish data for both the opportunity and necessity group. As stated in the previous section, this will give us an opportunity to check, at least at the national level, how the numerosity of these two groups behaved in probability when close to a "turning point ", close to the edge between a boom and a bust, regardless of the exogenous forces which provoked it as the states of the world are unobserved in the model. Observing our raw data and considering its time span, this was clearly possible for the past financial crisis as the event occurred in the middle of our time series. Let's look at Table 31 first, where the results for the opportunity groups are visible. A few points are in order: first of all, comparing the variance and autoregressive unrestricted models with the baseline one shows very similar transition probability results: everything else equal, the probability of staying in a low numerosity state from one time point to another is very high on average (from $96,5\%$ to $98,8\%$), while the probability to transition back from a high numerosity state to a low numerosity state is comparatively lower (from $50,7\%$ to $72,3\%$), showing signs of resilience and persistence across the boards for the genuine opportunity entrepreneurs. Furthermore, as we observe the models' state average, the average values appear to be, for the variance and autoregressive unrestricted models, pretty close in level across states: not only transitions are on average not very probable, but the average numerosity of the opportunity entrepreneur group appears to be quite constant across the two states. As we observe the baseline model we indeed see a bigger difference in the cross-state averages, where the high state one almost doubles the low state one. This is perhaps the only non-negligible difference, but according to the SBIC criteria we would however be more inclined to accept the results of the unrestricted models rather than the baseline one. In Figure 7, the predicted marginal, quarter by quarter transition probability of switching from the first state to the first state ("persistence probabilities ") are depicted against the numerosity of the opportunity entrepreneurs group. The picture refers to the probabilities predicted on the autoregressive switching model (column (III) in Table 7), selected according to the SBIC criterion, but all the models gave a similar picture. With the only exception of a decrease in the value given by the measurement change we explained in the data section, we observe how the probability of the low state have been stable across the whole analyzed period. This also includes the last financial crisis, highlighted inbetween two black solid vertical lines and ranging from the first quarter of 2008 to the first quarter of 2010. A similar conclusion can be drawn for the high state.

Let's now focus on the necessity entrepreneur group. Results for such group are visible in Table 32. Aside from a few statistically nonsignificant lags, the estimation for the necessity group is also plagued by far less precise probability estimates. In terms of unconditional average, the numerosity across states remained fairly similar. Selecting again the most relevant model in terms of the SBIC criterion and the significance of the probability estimates, we observe how both persistence in the low numerosity state and transitioning back to the low state form the high one appear fairly high (86% and $81,2\%$). In particular, the probability to go back to the low numerosity state, period by period, is generally lower by at least $8,9\%$ when $p_{2\longrightarrow 1}$ in column (II) of Table 32 is compared to its equivalents in Columns (I), (II) and (III) of Table 31, with the highest differ-

Table 31: ARMS(1/2) for opportunity entrepreneurs, Spain

	$Linear$	$MS - Variance$	$MS - Weights$
	(I)	(II)	(III)
$y_{t-1,s=1}$	0,493	0,877	0,870
	(0,105)	(0,039)	(0,037)
$y_{t-2,s=1}$	0,384		
	(0,104)		
$y_{t-1,s=2}$			-2,099
			(0,260)
$y_{t-2,s=2}$			
$\mu_{s=1}$	133929,700	116725,100	137193,700
	(17242,510)	(12591,060)	(4479,28)
$\mu_{s=2}$	249679,600	137413,300	139316,100
	(21292,310)	(18000,29)	(94868,432)
$\sigma^2_{s=1}$	18767,590	12647,180	13613,310
	(1478,316)	(1183,971)	(1091,362)
$\sigma^2_{s=2}$		94092,370	
		(36709,790)	
$p_{1\to1}$	0,988	0,965	0,966
	(0,012)	(0,026)	(0,024)
$p_{2\longrightarrow1}$	0,507	0,668	0,723
	(0,351)	(0,312)	(0,261)
$SBIC$	23,064	22,543	22,502

Markov application with lag and states for the linear (I), Variance switching (II) and Time weights switching (III) models. Standar errors in parenthesis.

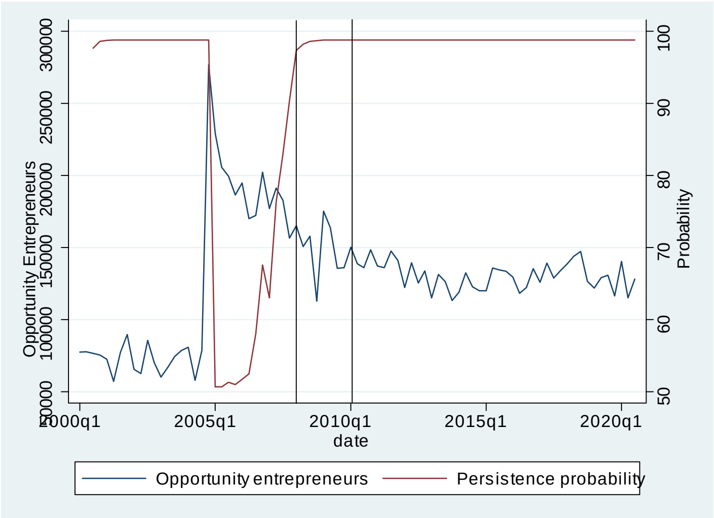

Figure 7: Marginal transition probabilities of the opportunity entrepreneurs group

Table 32: ARMS(1/2) for necessity entrepreneurs, Spain

	$Linear$	$MS - Variance$	$MS - Weights$
	(I)	(II)	(III)
$y_{t-1,s=1}$	0,626	0,368	0,356
	(0,129)	(0,096)	(0,107)
$y_{t-2,s=1}$	0,215	0,558	0,616
	(0,128)	(0,097)	(0,107)
$y_{t-1,s=2}$			0,048
			(0,193)
$y_{t-2,s=2}$			-0,390
			(0,174)
$\mu_{s=1}$	77890,500	71623,860	76515,460
	(8608,134)	(17558,660)	(4565,003)
$\mu_{s=2}$	71369,720	60813,680	89137,210
	(8556,276)	(15039,670)	(3627,535)
$\sigma^2_{s=1}$	11978,370	29934,550	7580,973
	(1038,043)	(11323,210)	(916,9579)
$\sigma^2_{s=2}$		8189,682	
		(869,348)	
$p_{1\to 1}$	0,000	0,329	0,860
	(0,002)	(0,275)	(0,084)
$p_{2\longrightarrow 1}$	0,882	0,062	0,812
	(0,131)	(0,050)	(0,220)
$SBIC$	22,099	21,771	21,663

Markov application with lag and states for the linear (I), Variance switching (II) and Time weights switching (III) models. Standar errors in parenthesis.

ence amounting to around $81,2\% - 50,7\% = 30,5\%$. **As a caveat, this is not enough to imply that the lowest persistence of the numerosity of the necessity entrepreneurs group and its highest probability of decreasing in value is connected in any way to the economic cycle (we did try to plot the marginal probabilities against the industrial production index, and found no straight evidence of that) but shows univocally that, not conditioned by the state of the economy, regardless of the nature of any given shock, an entrepreneur born out of necessity is always more likely to exit any given market when compared to the genuine entrepreneur who took an opportunity to enter it.** Patterns of persistence do exist in both groups, but are unavoidably different. The marginal probabilities of $p_{1\to 1}$ for column (III) in Table 32 are visible in Figure 8. As we have pointed out, the probability of low to low state transition stays fairly high across the board, save for some punctual spikes which however reflect a change in probability ranging from 81% to 86%. As stated before, the behavior of the GDP proxy does not appear to create systematic and sustained behavior in the marginal probabilities, but starting from 2012 the Figure reveals some degree of counterciclicality that should warrant further study.

10 Conclusions

We have found out that a form of stochastic persistence exists in the conditional average value of the opportunity and necessity entrepreneurs groups. The hysteresis phenomenon can as such be proved at the regional level once the heterogeneous effects of different economic structures across regions have been taken into account. Apart from time-wise dependence, we also tested for alternative forms of nonlinearity of the series. We could prove, at the aggregated panel level, that an alternative form of regime driven nonlinearity exists

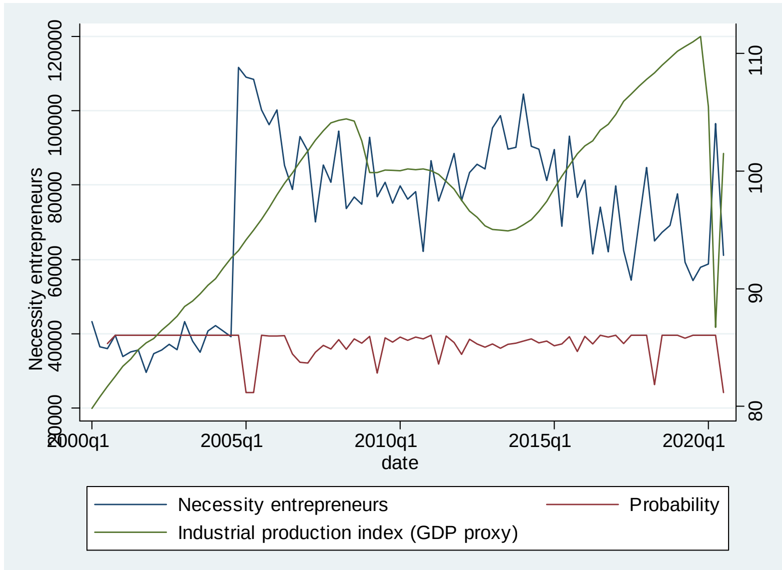

Figure 8: Marginal transition probabilities of the necessity entrepreneurs group plotted against numerosity and the industrial index

in both groups and can be framed by imposing unemployment as a weakly exogenous variable acting as a threshold between different regimes of convergence. As we looked for the best functional form, we have finally seen how multiple branch regressions models such as the threshold autoregression model describe asymmetries with a better fit than the linear ADF specification in a number of regional cases, while non negligible differences in terms of speed of adjustment can be found across economic sectors. This appears to be the case whether unemployment or GDP are chosen as weakly exogenous forces dominating the transition. The persistence of the opportunity and necessity groups is perhaps even more evident when a model with non-observed components are used as a mean to capture persistence at turning points (if there are any). Through a final Markov switching autoregressive application, we showed how both groups typically present persistence in level in two ideal states of the world, which are characterized by either a low or a high numerosity of entrepreneurs belonging to the two groups. Everything else equal, from time to time point, the probability of a necessity entrepreneur to exit the market will be consistently higher when compared to the same probability in the opportunity entrepreneur group. This appears to be the case regardless of any external force, as not even during the latest worldwide financial crisis the economic shock appears to have created a systematic, prolonged pattern in probability.

11 References

[1] C. Dawson, and A. Henley, “T’Push’ Versus ‘Pull’ Entrepreneurship: An Ambiguous Distinction?”, International Journal of Entrepreneurial Behaviour & Research, 18(6), pp. 697-71, 2012.

[2] R. Amit, and E. Muller, E.. "‘Push’ and ‘pull’ entrepreneurship.", Journal of Small Business and Entrepreneurship, 12(4), 64–80, 1995.

[3] E. Congregado, A. Golpe, and A. Van Stel, “The ‘recession-push’ hypothesis reconsidered”, International Entrepreneurship and Management Journal, 8(3), 2012, pp. 325-342.

[4] R.W. Fairlie, and F.M. Fossen, “Defining opportunity versus necessity entrepreneurship: two components of business creation,” NBER Working Paper No. 26377, October 2019, doi: 10.3386/w26377.

[5] S.W. Polachek, and K. Tatsiramos, "Change at Home, in the Labor Market, and On the Job ", Research in Labor Economics, Vol. 48, Emerald Publishing Limited, 2020, pp. 253-289. https://doi.org/10.1108/S0147-912120200000048008

[6] F.M. Fossen, “Self-employment over the business cycle in the USA: a decomposition”, Small Business Economics, 2020, pp. 1-19.

[7] F. Neymotin, “Necessity and Opportunity Entrepreneurship in Canada”, Review of Economic Analysis, 2020, 12(3).

[8] M. Caliendo, S. Künn and M.Weissenberger, "Catching up or lagging behind? The long-term business and innovation potential of subsidized start-ups out of unemployment ", Research Policy, 49(10), 2020.

[9] M. Caliendo, and A.S. Kritikos, "I Want to, But I also Need to": Start-Ups Resulting from Opportunity and Necessity”, IZA Discussion Paper No. 4661, 2009.

[10] T. Vogelsang, and P. Perron, "Nonstationarity hypothesis and level shifts with an application to purchasing power parity ", Journal of Business and Economic Statistics, 10, pp. 301-320, 1992

[11] J. Clemente, A. Montañes, M Reyes, "Testing for a unit root in variables with a double change in the mean ", Economics Letters, 59, pp. 175-182, 1998

[12] B. E. Hansen, "Sample splitting and threshold estimation", Econometrica, 68(3), pp. 575-603, 2000

[13] H. Tong, “Threshold Models in Non-linear Time Series Analysis. Springer New York”., 1983

[14] K. Hadri, "Testing for stationarity in heterogeneous panel data ", Econometrics Journal, 3, pp. 148-161, 2000

[15] K. S. Im, M. H. Pesaran, and Y. Shin., "Testing for unit roots in heterogeneous panels ", Journal of Econometrics, 115, pp. 53–74, 2003

[16] A. Gonzales, T. Terasvirta, D. Vand Dijk, and Y. Yang., "Panel Smooth Transition Regression Models ", Economic Letters, 59, pp. 175-182, 1998

[17] J. D. Hamilton, "A new approach to the economic analysis of nonstationary time series and the business cycle. ", Econometrica, 57, pp. 357–384,

1989, https://doi.org/10.2307/1912559.