Revista de Economía Aplicada E A *Número 62 (vol. XXI), 2013, págs. 61 a 90*

SELF-EMPLOYMENT: TRANSITION AND EARNINGS DIFFERENTIAL*

GEMA ÁLVAREZ
CARLOS GRADÍN
Mª SOLEDAD OTERO
Universidad de Vigo

In this paper we analyze the factors that influence transitions into self-employment in Spain using a discrete time duration model and, given the evidence of lower earnings among self-employees, we further explain the earnings differential between employees and self-employees using an Oaxaca-Blinder approach. The analysis is based on the European Community Household Panel (ECPH) for 1994-2001. According to our results, the factors explaining the transition into self-employment differ depending on the worker's previous status in the labor market. Additionally, we show that the observed earnings differential between self- and paid employees is a consequence of the selectivity bias into each labor status.

Key words: self-employment, longitudinal data, duration model, earnings differential.

JEL Classification: J82, J16, L26.

Entrepreneurial activity has been an important topic of research during the last few decades. Economists are concerned with self-employment because of the relationship thought to exist between entrepreneurship and economic development. For example, entrepreneurship is considered to play an important role in job creation, and one that provides an avenue out of poverty for many individuals. Also, new firms are thought to be involved significantly in innovative activity, promoting knowledge creation and fostering economic growth [see Audretsch (2007)]. All of this helps to explain why many studies have analyzed the determinants of the transition into self-employment in various countries, including Germany [see Georgellis and Wall (2005)], the U.S. [see Taniguchi (2002)], the UK [see Taylor (1996)] and Spain [see Congregado *et al*. (2006)].

The Spanish case appears to be of special interest because of the characteristics of its labor market, with lower employment rates, especially for women and young people, who are also affected by there being relatively fewer part-time jobs and more fixed-term contracts than in most developed countries. For example,

(*) Carlos Gradín acknowledges financial support from the Spanish Ministerio de Economía y Competitividad (Grant ECO2010-21668-C03-03/ECON) and Xunta de Galicia (Grant 10SEC300023PR).

data from 2011 indicate that Spain had a female employment rate of 55.5 percent, lower than that of Sweden (77.2%), Portugal (64.8%), France (64.7%) and Ireland (59.7%). The unemployment rate for Spanish women was 22.2 percent, higher than those of Sweden (7.5%), France (10.2%), Ireland (10.6%) and Portugal (13.2%). In Spain, the proportion of employees with a contract of limited duration was 25.3 percent, higher than those of Ireland (9.9%), France (15.2%), Sweden (16.4%) and Portugal (22.2%)[1]. All these features of the Spanish labor market may create the conditions for pushing those who are out of employment or whose jobs are precarious into self-employment. Additionally, the strong economic growth that the Spanish economy experienced after 1995 could have also contributed to employees entering self-employment in order to use their capital and skills to take advantage of new economic opportunities.

In this paper, we look at self-employment in Spain, focusing our attention on two important issues. First, we analyze entry into self-employment, introducing two relevant methodological points that have not received much attention so far in this field. Firstly, we take into account the potential bias of unobserved heterogeneity. When it comes to considering the probability of becoming self-employed from an empirical point of view, there may be unobserved factors that differ among individuals, such as the ability to start up a business or the preference for leisure or for a flexible job, which could have an impact on the probability of becoming self-employed. As pointed out in Bover, Arellano and Bentolila (1996) and Carrasco (2002), ignoring this unobserved heterogeneity may lead to biased estimates. Secondly, we use lagged explanatory variables to explain entry into self-employment. In this way, as Georgellis and Wall (2005) point out, the problem of the explanatory variables being consequences rather than causes of self-employment can be avoided.

The fact that certain individual attributes are associated with a higher probability of being self-employed implies that employees and the self-employed differ in their characteristics, and so may also differ in earnings. This raises the second and complementary issue addressed in this paper: the determinants of the earnings differential between the self-employed and employees. Given that Spanish self-employed workers obtained far less earnings than employees (despite working more hours), using a Oaxaca-Blinder approach and after controlling for selectivity bias, we analyze the determinants of these earnings differentials, considering especially to what extent they are explained by different productivity-related characteristics of workers or, alternatively, by different returns to such attributes. This is a topic that has received very little attention in the literature so far, in contrast with the wage gaps among employees by gender, ethnicity, nationality, and sector (private or public), which have been widely discussed. A detailed decomposition of the earnings differential makes it possible to identify which specific attributes or returns contribute the most to explaining the differential in earnings.

(1) Source: http://epp.eurostat.ec.europa.eu/portal/page/portal/employment_unemployment_lfs/data/main_tables.

The paper is organized as follows. Section 1 reviews the empirical literature related to the two topics, while Section 2 describes the data and presents the models used and the estimation procedures. Section 3 shows the empirical evidence found and, finally, Section 4 summarizes the main conclusions.

1. Empirical literature

1.1. Entry into self-employment

There is a vast literature that has examined entry into self-employment from an empirical point of view. In order to explain self-employment decisions, a number of demographic, economic and labor characteristics of self-employed individuals have been considered.

Among the demographic characteristics of individuals, the most common are education, age, gender, marital status, and the presence of small children in the household. Although the increase in self-employment over recent decades has been greater among women [see Hughes (2003)], entrepreneurs are still mostly men. Accordingly, empirical research tends to find a negative relationship between being female and self-employment [see Arenius and Minniti (2005), Hammarstedt (2004), and Blanchflower and Oswald (1998)]. As for the other variables, the literature provides no clear results about their relationship with the probability of entering self-employment. For example, education is expected to be positively related to the probability of becoming self-employed because more highly-skilled people may have more information about business opportunities and a greater ability to act on them [see Rees and Shah (1986)]. This is supported by Blanchflower and Meyer (1994), and Fujii and Hawley (1991). However, a higher level of education may also involve higher salaries, adversely affecting the probability of becoming self-employed. This negative effect has been found, for example, by Kidd (1993), De Wit and Winden (1989), and Evans (1989). By age, the probability of entering self-employment is expected to be higher among younger people because they have more market opportunities and are more prone to assume the risks involved in a business [see Blanchflower and Meyer (1994)]. However, older people are more likely to set up a business if they face more difficulties in finding a paid job and, at the same time, have accumulated the necessary financial resources. Similarly, the stability of marriage can provide an appropriate framework to take the risk involved in self-employment [see Le (1999)], but family responsibilities, especially in the presence of children, may work in the opposite direction. Even in this case, the greater flexibility of self-employment could help married people with children to balance work and family life.

Economic determinants, such as wealth and the conditions of the labor market, may also affect the probability of becoming self-employed. Wealth is expected to have a positive and significant impact on the probability of starting a business because doing so often requires a significant initial investment and there are capital and liquidity constraints [see Dunn and Holtz-Eakin (2000), Johansson (2000), Blanchflower and Oswald (1998), Taylor (1996), Evans and Jovanovic (1989), and Evans and Leighton (1989)]. The conditions of the labor market, measured by the rate of unemployment, could, a priori, either increase or reduce the rate of

self-employment [see Tervo (2006)]. Its effect on self-employment is often interpreted in terms of the so-called 'pull' and 'push' factors. 'Pull' factors are stronger when conditions are good, due to the better prospects for a business and the higher probability of finding a paid job if the business fails [see Carrasco (1999)]. On the other hand, less favorable market conditions, with high unemployment rates, can lead to a 'push' effect in which one starts a business as the only way out of unemployment or inactivity. Empirical evidence has supported both factors: 'pull' [see Blanchflower and Oswald (1991), and Taylor (1996)] and 'push' [see Evans and Leighton (1989)].

Finally, the characteristics of the job, such as tenure, type of contract, working time and size of the firm, can also be determinants of the transition from paid employment to self-employment. The effect of job tenure is expected to be negative because both the accumulation of specific human capital and exit costs increase with seniority [see Carrasco and Ejrnaes (2003)]. However, the accumulation of experience, capabilities, skills, and assets promoting entry into self-employment increase with seniority, too. Authors such as Carr (1996) and Devine (1994) have shown a positive relation between tenure and self-employment. Further, fixed-term and/or part-time employees are expected to have a higher probability of becoming self-employed because of their lower opportunity costs. The size of the employee's firm is usually found to have a negative correlation with his or her probability of entry into self-employment [see Blanchflower and Meyer (1994)]. This effect may be explained by the non-pecuniary benefits offered by larger firms, such as greater job security. Similarly, working in the public sector also raises the opportunity costs of self-employment due to greater job stability [see Leoni and Falk (2010), and Carrasco and Ejrnaes (2003)].

1.2. The earnings differential between salaried employees and self-employees

Despite the large earnings differentials that can exist between self- and paid employees, of which Spain is a clear example, the sources of such gaps remain unexplained, in contrast with the abundant evidence about wage gaps among employees based on gender, sector or ethnicity. Some authors have studied particular aspects of self-employment earnings. For example, Hundley (2001) and Tansel (2002) examined the gender earnings gap among self-employees, while Moore (1983) compared the gender and race earnings gaps among the self-employed with those among paid employees as a measure of discrimination. But none of these authors aimed to explain the earnings gap between self- and paid employees or to what extent this gap can be attributed to the different human capital endowments of the two groups. This asymmetry is obviously related to the lack of adequate information about self-employees' earnings in most surveys, which are usually collected only on an annual basis and tend to be more underreported than earnings from paid employment. However, we think that we will not be able to understand the functioning of self-employment in a country with large earnings differentials if we ignore the sources of such gaps.

The literature on earnings differentials is generally based on the estimation of Mincerian regressions in which the wage (in logs) is a function of the worker's productivity (mainly captured by individual attributes, such as education and experience), several characteristics of the job (such as industry and occupation), and a few

control variables to take into account the economic cycle (year) or regional disparities in wages. Regressions are usually estimated separately for each gender, in order to account for the different payment schemes for men and women in the labor market. The main question addressed by the literature of wage differentials is, after controlling for any selection bias, to what extent the observed differential can be explained by differences in productivity-related endowments in both groups of interest. This is usually referred to as the 'characteristics effect', with the remaining unexplained differential being the consequence of attributes having a different impact on earnings in both groups, which is called the 'coefficients effect'[2]. These effects can be estimated using the Oaxaca (1973) and Blinder (1973) approach.

2. Data, models and procedures

2.1. Data

The data used in our empirical analysis come from the European Community Household Panel (ECHP) for Spain. The longitudinal design of the ECHP made it possible to follow up and interview the same set of individuals over eight consecutive years: 1994-2001. ECHP data is collected by National Data Collecting Units in collaboration with the Statistical Office of the European Communities, Eurostat. The annual data provided in these surveys contains information, detailed and homogenized, on personal and family characteristics, as well as on the labor history of the individuals.

We restricted the analysis to people between 19 and 55 years old in the first wave and who are either working (15 or more hours a week) in paid employment or out of employment (unemployed or inactive). We look at those working 15 or more hours a week because key variables for the analysis (e.g., self-employment status, part-time work or firm size) are known only for this group of workers. Finally, those working in the agricultural sector in any wave were excluded.

2.2. Discrete time duration model

In this section, we present the discrete time duration model used in the analysis of the transition times from non-working or paid employment to self-employment. As usual, the model will specify the impact of the individual's characteristics on the hazard rate; that is, on the instantaneous probability of moving to the self-employment state. More explicitly, if T_i is the transition time of the *ith* individual and if h_{it} denotes his/her hazard at time t, we will have $h_{it}(X_{it}) = P(T_i = t | T_i \geq t, X_{it})$, where X_{it} is the vector of time-varying covariates. A commonly used model is Cox's proportional hazards model [see Cox (1972)]:

$$h_{it}(X_{it}) = h_{0t} exp(\beta X_{it}), \qquad [1]$$

where β is a vector of parameters and h_{0t} is the baseline hazard at time t. The popularity of the Cox model is due to (a) the immediate interpretation of $exp(\beta_j)$ as a

(2) See, for example, Cain (1986) for a classic survey that offers a detailed reference to the most important theories that attempt to explain wage differentials based on Mincer (1974) models.

proportionality risk factor associated with a unitary increase of the *jth* explanatory variable, (b) the fact that it can be estimated without specifying the baseline hazard (the non-parametric part of the model), and (c) the fact that the model can be applied in the presence of censored observations (a common issue in duration analysis), and also allows for time-varying covariates.

Expression [1] refers to a continuous transition time T_i. In our case, time is rounded to years (from 1 to 7), so we used the discrete version of the Cox model, which corresponds to the following specification [see Kalbfleisch and Prentice (1980)]:

$$log\ (1 - h_{it}\,(X_{it})) = exp\ (\beta X_{it})\ log\ (1 - h_{0t}). \qquad [2]$$

This is the well-known clog log model (from complementary log-log). Jenkins (1995) suggested the introduction of dummies for each year t from the panel in the vector X_{it} as a flexible (semi-parametric) version of the model. Thus, in equation [2], $log\ (1 - h_{0t}) = -exp\ (\gamma_t)$, with $t = 1,.., 7$, and where $_t$ are the parameters measuring the duration dependence of the model.

The clog log model [2] was estimated using the maximum likelihood principle. Following Jenkins (1995), the log-likelihood function is:

$$logL\ (\beta,\ \gamma) = \Sigma_i\ \Sigma_t\ y_{it}\ log\ [h_{it} / (1 - h_{it})] + \Sigma_i\ \Sigma_t\ log\ (1 - h_{it}), \qquad [3]$$

where $h_{it} = h_{it}\ (X_{it})$ for simplicity of notation, and where y_{it} is equal to 1 if the *ith* individual transitions at time t, and *zero* otherwise. The sum Σ_t is taken over the periods (years) of the panel in which the individual is observed. Once the β parameters are estimated, they can be interpreted in the continuous subjacent model [1]. This likelihood function has been modified by Meyer (1990) to account for gamma distributed unobserved heterogeneity.

2.3. Decomposition of the differential in annual earnings

In order to explain the reasons for the large differential in average annual earnings between employees and self-employees, we use the well-known regression-based Oaxaca-Blinder decomposition. We estimate one Mincerian equation of the annual earnings (in logarithms) for each sector (self-employee: j = 1, and employee: j = 2), which, omitting time subscripts for simplicity, can be expressed as[3]:

$$\ln\left(y_i^j\right) = X_i^j \beta^j + u_i^j \qquad [4]$$

where X_i^j is the corresponding vector of the worker's human capital characteristics (age, education, tenure, previous experience and unemployment spells) and other controls for the nature of the job (such as industry and occupation), as well as re-

(3) Even though the best option would be to consider the logarithm of hourly earnings, it is not possible to reasonably estimate the number of hours worked by self-employees in the ECHP. This is because earnings from self-employment are annual and workers only declare the current number of hours worked in a week, with no information about how many weeks or months they worked. This is not the case for employees, as they declare their monthly wage too. An approximation of hourly earnings, assuming that all workers worked during the entire year, gives similar results to those reported in this paper.

gional and time dummies[4]. β^j is the associated vector of OLS coefficient parameters and the error term. The observed average differential between earnings in the two sectors can be rewritten as the sum of two terms:

$$\overline{\ln(y^2)} - \overline{\ln(y^1)} = \bar{X}^2\hat{\beta}^2 - \bar{X}^1\hat{\beta}^1 = \left(\bar{X}^2 - \bar{X}^1\right)\hat{\beta}^2 + \bar{X}^1\left(\hat{\beta}^2 - \hat{\beta}^1\right) \quad [5]$$

The first term on the right hand side of the last equation is the aggregate characteristics effect; that is, the earnings gap that can be explained by the difference in endowments valued at the returns for employees' characteristics. The second term is the aggregate coefficients effect; this is the gap that can be attributed to different returns for worker's characteristics in both sectors, evaluated at self-employees' characteristics. This last term is by construction the unexplained part of the earnings differential and indicates to what extent a given worker's attribute (for example, attained education) has a different impact on earnings depending on the sector in which he or she works.

The estimates in the OLS regression might be biased and inconsistent due to self-selection of individuals into either non-work, self-employment or paid work (respectively, $j = 0, 1, 2$). In order to take this into account, we use the technique of Lee (1983) and Trost and Lee (1984), which extends the well-known approach of Heckman (1974) for dealing with sample selection bias to the case of more than two alternative outcomes. This solution consists of first estimating the probability of choosing alternative *j*, using a conditional multinomial logit model and taking people out of work as the reference group:

$$P^j = \exp(Z\hat{\alpha}^j) \Big/ \left(1 + \sum_{j=1}^{2} \exp(Z\hat{\alpha}^j)\right) \quad [6]$$

where Z is the vector of explanatory variables affecting sectorial choice such as worker's age, marital status, gender, education and capital/property income, and additional information on the household that could influence the worker's decision, such as the number of children under six and the amount of income from other members, as well as regional and time dummies as controls. $\hat{\alpha}^j$ is the vector of parameter estimates for the probability of choosing alternative *j*. Based on these probabilities, we can construct the selection term for each alternative *j,* as follows:

$$\lambda^j = \phi\left(H^j\right) \Big/ \Phi\left(H^j\right), \qquad where \qquad H^j = \Phi^{-1}\left(P^j\right) \quad [7]$$

with ϕ and Φ being, respectively, the standard normal density and distribution functions. Finally, this term is included in the earnings regression equation as an explanatory variable, with $\hat{\alpha}^j$ being the corresponding coefficient:

$$\ln\left(y_i^j\right) = X_i^j\beta^j + \theta^j\lambda^j + u_i^j \quad [8]$$

The analysis of the sectorial differential in earnings is undertaken separately for men and women, given that returns to characteristics are known to vary across gen-

(4) Using the panel structure of ECHP, we link characteristics of each year with annual earnings of the same year (but declared in the following wave of the survey).

ders. Further, this allows us to use the same methodology to explain the large earnings gap by gender within each sector, using the coefficients of men as the reference.

$$W^k_{\Delta\beta}$$

Finally, in order to evaluate the individual contribution of each variable (or set of variables) to the total explained differential, known as the detailed decomposition, we estimate a set of individual contributions of characteristic k (k = 1, …, K) to the aggregate characteristics effect, $W^k_{\Delta x}$, and the coefficients effect, $W^k_{\Delta\beta}$, such that:

$$W^k_{\Delta\beta} = \bar{X}^0_k(\beta^1_k - \beta^2_k) \quad [9]$$

However, an additional and well-known problem which needs to be addressed is that detailed decompositions of coefficients effects suffer from an identification problem. This is because the contribution of a dummy variable to the coefficients effect will vary with the choice of the reference group, and this applies to any set of dummy variables. To tackle this difficulty, we use normalized regressions in computing the detailed effects, as proposed by Suits (1984), Gardeazabal and Ugidos (2004) and Yun (2005a, 2005b). This method has the advantage of being invariant with respect to the "left-out" reference category in computing the contribution of dummy variables to the coefficients effect. Further, it alters neither the detailed characteristics effect nor the contribution of continuous variables to the coefficients effects that are unaffected by the identification problem.

3. Empirical results

In this section, we present the results of our empirical analysis on entry into self-employment as well as on the earnings differential between employees and self-employees.

3.1. Transition into self-employment

3.1.1. Variables and descriptive analysis

We analyzed transition into self-employment during the period 1995-2001[5]. The first transition into self-employment considered was that produced in 1995 (instead of in 1994) because we used lagged explanatory variables (i.e. we used characteristics in 1994 to explain the transition in 1995). This ensures that explanatory variables are consequences rather than causes of self-employment [see Georgellis and Wall (2005)]. The final sample was composed of 3,679 workers and 4,503 non-workers who were followed up until the first transition into self-employment occurred (non-censored observations) or until the individual was no longer observed or transitioned to another state of employment (censored observations).

The explanatory variables used in the estimation are based on our review of the literature (see Section 1.1. above). These variables refer to demographic and economic characteristics of the individual and his or her family, employment history,

(5) In defining self-employees, we discarded those who were unpaid workers in a family business.

and other control variables. Among the demographic characteristics, we take into account the educational level attained by each individual, age, gender, marital status, and if there were children in household. Among the economic and labor characteristics, we included a measure of household wealth, the unemployment rate by demographic group, and an indicator of whether the individual had worked before. This last variable was included in order to take into account differences in labor market opportunities due to divergent previous labor experience. All estimations include regional and time dummies as additional control variables. In Table A.1 in the appendix, we report information about the construction of all these variables.

Table I shows the sample means calculated for workers and non-workers in the first period they were observed. The last column in the table shows the *p*-value associated with a test on the equality of means between groups.

Table I: Sample means by initial labor status

Variable	Workers	non-workers	P>|Z|
Change to self-employment	0.043	0.054	0.015
Age			
19 to 30	0.295	0.451	0.000
31 to 45	0.500	0.333	0.000
>45	0.205	0.216	0.263
Education			
Less than secondary	0.500	0.622	0.000
Secondary	0.200	0.249	0.000
College	0.300	0.130	0.000
Gender			
Woman	0.351	0.700	0.000
Marital Status			
Married	0.653	0.561	0.000
Married woman	0.193	0.471	0.000
Married man	0.460	0.090	0.000
Other variables			
Children under 6	0.268	0.240	0.017
Wealth	41.338	41.230	0.976
Unemployment rate	22.9	31.1	0.000
Experience	0.503	0.640	0.000
No. of observations	3,679	4,503	

Source: Our estimations using ECHP.

Those who are initially out of work are more likely to become self-employees in Spain than those who already have a job. More specifically, 4.3 percent of workers become self-employees over the sample period compared to 5.4 percent of non-workers. This figure rises to 6.7 percent if we look at unemployed people. This evidence supports the view that self-employment is the only alternative for some people to become employed, due to constraints in the labor market. Regarding the characteristics of both samples, non-workers tended to be younger than workers, who were more likely to have a college degree. With respect to gender, the worker sample included a large proportion of men while the non-workers group included relatively more women. Looking at marital status by gender, it can be seen that there are considerably more married women among non-workers (47.1 percent compared to 19.3 percent) and more married men among workers (46.0 percent compared to 9.0 percent). This is obviously related to the traditionally lower employment rate of women, especially married women, and of young people in Spain, compared with other developed countries. Both groups were similarly wealthy on average. Finally, note that non-workers faced more unfavorable market conditions (higher unemployment rates for their demographic groups), while they also tended to have greater previous labor experience.

In the analysis of the transition of employees into self-employment, we also took into account a number of characteristics of the job, such as tenure, type of contract, working time, size of the firm, and whether the worker was employed in the public sector. This last attribute was introduced into our estimations in relation to educational level, in order to take into account the fact that the opportunity cost of entering self-employment may be higher among those working in the public sector, and this could differ across educational levels. The construction of these variables is also described in more detail in Table A1 in the Appendix. In Table II, we summarize the characteristics of jobs for the sample of workers. The means have been calculated separately for employees who changed to self-employment during the sample period (first column) and those who did not (second column). Again, the last column in the table shows the *p*-value of testing the equality of means between groups. As expected from the previous discussion, workers who become self-employed are likely to have less tenure and more likely to work in a part-time or fixed-term job, as well as in smaller firms. Finally, it can be seen that the higher the educational level of public sector employees, the lower their likelihood of becoming self-employed.

We start the study of transition into self-employment with a descriptive analysis based on the estimation of Kaplan-Meier survival functions[6]. In Table III, we report the log-rank test for the comparison of several groups of individuals, attending to their specific characteristics. Here, the null hypothesis is that the survival function is the same among the groups being compared. Hence, groups of individuals for which the null is not rejected correspond to groups exhibiting a similar pattern when moving to self-employment over time.

(6) Kaplan and Meier (1958).

Table II: Sample job characteristics for workers group

Variable	Change	No change	P>|Z|
Tenure	0.248	0.341	0.016
Part-time	0.102	0.066	0.081
Fixed-term contract	0.310	0.227	0.069
Large firm	0.147	0.278	0.000
Education interacting with public sector			
Less than secondary	0.019	0.077	0.007
Secondary	0.013	0.056	0.019
College	0.032	0.158	0.000
No. of observations	157	3,522	

Source: Our estimations using ECHP.

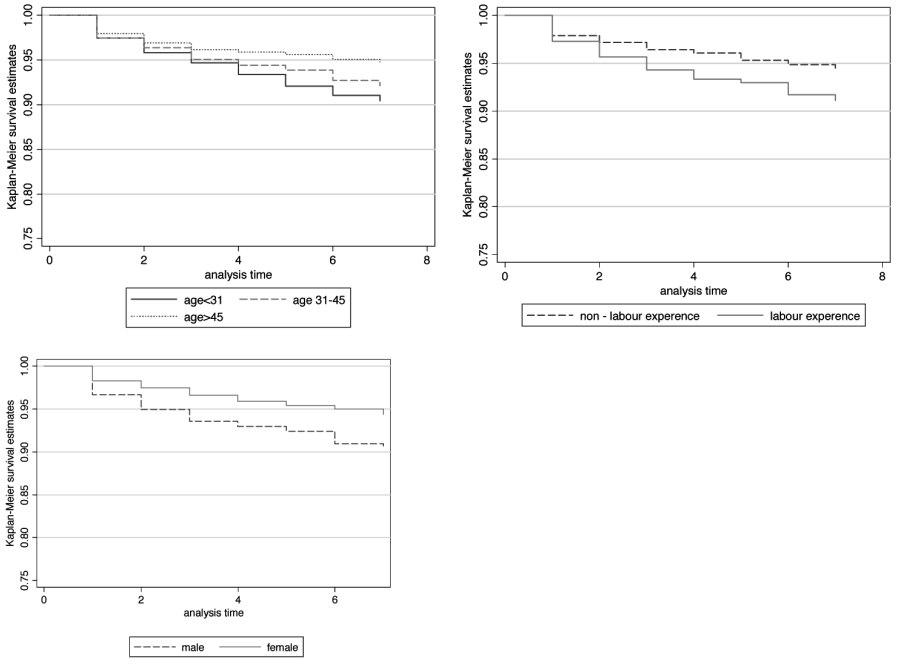
From Table III, we see that most groups under comparison have a different survival function, thus revealing the influence on transition time of gender, labor experience, age, job tenure time, public/private sector of activity, type of contract (permanent/non-permanent), and firm size. However, no differences were found (at a 5 percent significance level) when comparing groups by education, marital status, presence/absence of children, activity in industry/services, or part-time/full-time contract.

Table III: Log-Rank test for Kaplan-Meier survival functions

	Statistic	P>|Z|
Man/Woman	35.13	0.00
Non-labor experience/labor experience	18.16	0.00
Primary /secondary /collage education	0.13	0.93
Married /single / other marital status	0.97	0.61
Age <31 / age 31- 45 / age >45	11.23	0.00
Children under six	0.14	0.71
Job tenure >15 year / job tenure<15 year	13.32	0.00
Industry/ services	1.88	0.17
Public sector/ private sector	41.29	0.00
Fixed - term contract / permanent employment	13.46	0.00
Part-time job /full-time job	2.08	0.14
Large firm (50+ workers) / small firm (<50 workers)	43.95	0.00

Source: Our estimations using ECHP.

Figure 1: KAPLAN-MEIER SURVIVAL FUNCTIONS BY AGE, GENDER AND LABOR EXPERIENCE

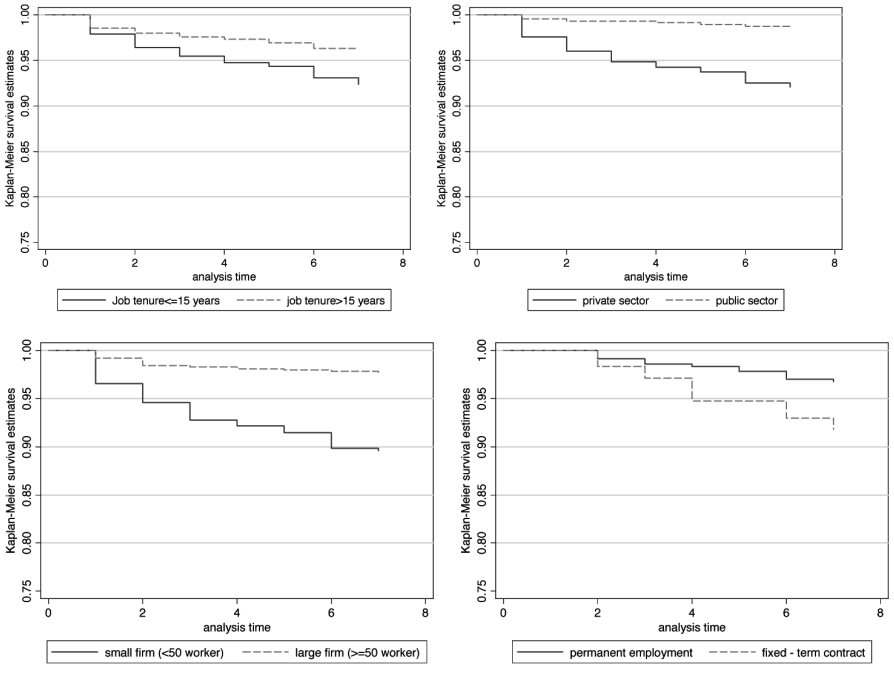


Source: Our estimation using ECHP.

Figures 1 and 2 display the Kaplan-Meier survival curves for the groups that show a different transition pattern. The first figure displays the survival functions for all individuals (both workers and non-workers) according to their age, sex and labor experience. The second focuses on job characteristics of workers, such as job tenure, firm size, sector (private or public), and type of contract (permanent or fixed-term).

One general fact that can be seen in Figures 1 and 2 is that survival decreased slowly during the follow-up for all groups, thus meaning a relatively small rate of movement to self-employment. After seven years in the same state, either employed or out of employment, only about 10 percent of the original sample had transitioned into self-employment. The profile of the type of individual most likely to change his or her initial status is that of a young male (below 30 years old), with labor experience, working with a fixed-term contract in a small firm in the private sector, with 15 years or less of job tenure.

Figure 2: KAPLAN-MEIER SURVIVAL FUNCTIONS BY JOB TENURE, SECTOR, FIRM SIZE AND TYPE OF CONTRACT



Source: Our estimation using ECHP.

The largest differences among the compared groups were found when changing from a small firm to a large firm, the former group showing a seven-year survival rate of 0.08 points below the latter. Differences were smaller for the other comparisons, ranging from 0.03 in the case of labor experience and job tenure to 0.05 in the case of private vs. public sector.

3.1.2. Duration model

With regard to the parametric estimations, the results of the proportional hazard model are reported in Tables IV and V[7]. In the first case, we estimate a model for the whole sample including three dummy variables that indicate if the individual was employed (omitted), unemployed or inactive in the first period. In this way we are able to test for the existence of differences across groups in their likelihood of becoming self-employed. Secondly, Table V shows estimated models for two sub-

(7) Estimates were obtained using STATA 11.0 the pgmhaz8 dofile created by S. Jenkins.

samples: one consisting of those employed in the first period (workers), and another integrated by those out of work in the first period (non-workers). The estimated coefficient, the hazard ratio, and the corresponding *p*-value are displayed in each case.

Our first interest was to test whether unobserved heterogeneity is a potential source of bias in estimating the determinants of transition into self-employment. According to our estimates of the LR test, unobserved heterogeneity was not significant in either sample. That is, unobserved sources of heterogeneity across individuals that could make them more prone to start a business, such as differences in their preferences or abilities, do not contribute significantly to explaining transitions into self-employment in Spain. Thus, we could focus the analysis on the role played by observed characteristics.

Table IV: TRANSITIONS INTO SELF-EMPLOYMENT (DISCRETE DURATION MODELS)

	All (workers and non-workers)		
	Coeff.	Hazard Ratio	P>|Z|
Unemployment	1.01	2.75	0.00
Inactive	0.73	2.08	0.00
Secondary education	-0.03	0.97	0.81
College	0.06	1.06	0.65
Age 31-45	-0.17	0.84	0.23
Age >45	-0.59	0.55	0.00
Woman	-0.71	0.49	0.00
Married	0.17	1.19	0.23
Children under six	-0.05	0.95	0.61
Wealth	0.03	1.03	0.80
Unemployment rate	-0.01	0.99	0.06
Experience	0.27	1.31	0.02
Intercept	-3.60	0.03	0.00
LR test (gamma var. = 0)			
Prob.> = chibar2		0.5	
No. of observations		29,437	

Note: Time and regional dummies have been used in all regressions. Reference category: Employed, 30 years old or younger single man, with primary education, without previous labor experience. Wealth appears multiplied by 10^6.

Source: Our estimations using ECHP.

Estimates in Table IV point to the unemployed as those who are the most likely to enter into self-employment. More specifically, the hazard ratio indicates that being unemployed in the first period increases by 175 percent the probability

Table V: Transitions into self-employment (discrete duration models)

	Non-workers Prob(SE_t/NW_{t-1})			Workers Prob(SE_t/W_{t-1})					
	Coeff.	Hazard Ratio	P>|Z|	Coeff.	Hazard Ratio	P>|Z|	Coeff.	Hazard Ratio	P>|Z|
Unemployment	0.34	1.40	0.04	–	–	–	–	–	–
Secondary education	0.00	1.00	1.00	0.05	1.05	0.81	0.45	1.57	0.03
College	0.59	1.80	0.00	-0.38	0.68	0.07	0.24	1.27	0.28
Age 31-45	0.11	1.12	0.58	-0.35	0.70	0.11	-0.11	0.90	0.63
Age >45	-0.41	0.66	0.11	-0.61	0.54	0.04	-0.35	0.70	0.24
Woman	-0.97	0.38	0.00	-0.69	0.50	0.01	-0.75	0.47	0.00
Married	0.46	1.58	0.02	0.03	1.03	0.90	0.12	1.13	0.59
Children under six	-0.06	0.94	0.64	0.05	1.05	0.78	0.05	1.05	0.76
Wealth	-0.35	0.70	0.33	0.65	1.91	0.00	0.62	1.86	0.00
Unemployment rate	-0.01	0.99	0.23	0.00	1.00	0.90	0.00	1.00	0.76
Experience	-0.04	0.96	0.80	0.50	1.65	0.00	0.11	1.12	0.62
Secondary education*public sector	–	–	–	–	–	–	-2.04	0.13	0.00
College*public sector	–	–	–	–	–	–	-1.62	0.20	0.00
Tenure	–	–	–	–	–	–	-0.13	0.88	0.28
Part-time job	–	–	–	–	–	–	0.49	1.63	0.13
Fixed-term-contract	–	–	–	–	–	–	0.24	1.27	0.36
Large firm	–	–	–	–	–	–	-1.27	0.28	0.00
Intercept	-3.03	0.05	0.00	-3.72	0.02	0.00	-2.97	0.05	0.00
LR test (gamma var. = 0) Prob.> = chibar2		0.5			0.5			0.5	
No. of observations		14,409				15,028			

Note: time and regional dummies have been used in all regressions. Reference category: Inactive, 30 years old or younger single man, with primary education, without previous labor experience. Wealth appears multiplied by 10^6.

Source: our estimations using ECHP.

of becoming self-employed, compared with being employed which is the reference category. For inactive individuals the probability increase is 108 percent. Therefore, the unemployed are the most likely to enter into self-employment, followed by the inactive. On the contrary, the employed are those with the lowest transition probability. These results are in line with the argument that self-employment could be the only alternative for some people to become employed.

With respect to the other variables, global estimates do not reflect a significant effect of education, while older individuals have a lower probability of entering self-employment. As expected, estimates indicate that women are less prone to enter into self-employment than men. More precisely, the probability of a woman becoming self-employed, ceteris paribus, is about half that of a man. By contrast, marital status and children do not significantly influence transition into self-employment. In order to analyze if the effect of marital status on the probability of being self-employed varies across gender [see Taniguchi (2002), and Carr (1996)], we estimated the model including interactions between marital status and gender variables[8]. Results suggest that married men are more likely to become self-employed than unmarried men. This result is in line with the argument that the stability of marriage can provide an appropriate framework to take the risk involved in self-employment [see Le (1999)]. For women, however, the probability of being self-employed does not depend on their marital status. Finally, while the unemployment rate reduces the probability of becoming self-employed, previous labor experience increases it.

Separate models for workers and non-workers in the initial wave are collected in Table V. Among non-workers, the unemployed are still the most likely to become self-employed. In this case, being unemployed increases by 40 percent the probability of becoming self-employed compared with being inactive. In contrast to global estimates, estimated coefficients for non-workers show that a higher level of education has a positive effect on the probability of becoming self-employed. Indeed, if we look at the hazard ratio by education, we observe that holding a college degree increased by 80 percent the probability of entry into self-employment, compared with having achieved only primary education (which is the reference case in the regression). A positive effect of education on the probability of entering self-employment for those out of employment has also been shown, for example, by Aguado, Congregado and Millán (2002), also for Spain[9], and by Georgellis and Wall (2005) for Germany. From a theoretical point of view, this is consistent with the hypothesis that education improves the information an individual has about business opportunities as well as the required abilities and skills. On the other hand, no significant differences in the likelihood of entry into self-employment were observed by age.

Like in the case of global estimates, non-working women are also less prone to enter into self-employment, which is consistent with the previous literature. In particular, non-working women are 62 percent less likely to become self-employed

(8) Estimates are not shown to save space but are available upon request.

(9) These authors, using the ECHP for a shorter period, estimated standard logit regressions without any dynamic structure.

than non-working men. By contrast, married individuals are 58 percent more likely to become self-employed than unmarried individuals. However, if we estimate the model including marital status interacted with gender, we find again that this positive effect occurs only among men. According to our estimations, the presence of children does not seem to be affecting the probability of those out of employment entering into self-employment. Again, this result is in line with the evidence shown by Georgellis and Wall (2005), and Aguado, Congreado and Millán (2002). Finally, other economic and labor variables, such as wealth, unemployment rate, and having had previous labor experience, do not affect the probability.

Regarding the workers group, Table V provides two different sets of estimates: one that includes the same demographic and economic variables considered for non-workers, allowing direct comparison of both groups, and another that also includes job-specific characteristics in order to analyze the role of previous labor market status in explaining the likelihood of a transition.

Interestingly, the effect of a worker's education and age depends on the specification of the model. Holding a college degree reduces the probability of entry into self-employment by 32 percent compared with primary education when the job's characteristics are not taken into account. This result is the opposite of the one found for non-workers, but has been shown before in the literature [see Hammarstedt (2004), Kidd (1993), De Wit and Winden (1989), and Evans (1989)]. This negative effect is consistent with the hypothesis that a higher level of education may involve higher salaries, which may in turn be negatively related to the probability of becoming self-employed. Indeed, educated workers are more likely to become employed in high-wage occupations and have the greatest possibilities of being promoted; hence, self-employment may be less desirable for individuals with higher education, provided they are inserted into the labor market. In line with these arguments, it can be observed that the negative effect of education completely vanishes when the job's characteristics are included in the model, especially when education dummies interact with a dummy indicating whether the individual was employed in the public sector. It could be that this negative effect of higher education is driven by the fact that many skilled workers have a stable job in the public sector and, thus, a lower probability of becoming self-employed (a significant negative effect). Similarly, the probability of becoming self-employed declined significantly with age in the first estimation: the probability of entry for individuals over 45 years old was 46 percent lower than that for those aged 30 or less. Again, this age effect vanishes when job variables are introduced into the model because it is driven by the specific characteristics of the jobs held by most young people. For example, young people in Spain are more likely to work with fixed-term contracts and obviously have less tenure in their jobs, and both these characteristics increase the probability of transitioning into self-employment, even if with low significance.

As in the case of non-workers, women have a probability of setting up a business that is about half that of men (50-53 percent lower, depending on the model specification). By contrast, marital status does not affect the probability of becoming self-employed among workers once gender has been taken into account. Further, neither the presence of children under six, nor the unemployment

rate seem to affect the probability of becoming self-employed for workers, as they do for non-workers. Nevertheless, higher wealth is a very significant variable for explaining the probability of workers moving from paid to self-employment, as expected, while this variable was not found to be significant for non-workers. As noted before, a positive relation between wealth and self-employment is common in this field [see Dunn and Holtz-Eakin (2000), Johansson (2000), Blanchflower and Oswald (19989, Taylor (19969, Evans and Jovanovic (1989), and Evans and Leighton (1989)]. Theoretically, the less critical his or her restrictions on capital, the greater the probability that an individual will choose to enter self-employment. Our results indicate that restrictions on capital are a determinant of the probability of starting a business among employees but not among non-workers. A plausible hypothesis for this result is that the businesses they start up involve different amounts of initial investment.

Estimates excluding job-specific characteristics show that the probability of transition for those who have worked before is 65 percent higher than for those who have not. Nevertheless, this variable may be correlated with some of the job characteristics considered. For example, it is expected that those who have worked before have less seniority in their jobs or do not work in the public sector. Thus, when these features are taken into account, the positive effect becomes statistically insignificant. Finally, and also in line with theory, the probability of entry into self-employment is lower among paid employees of larger firms. Indeed, working in a firm with 50 or more employees reduced the probability of moving from paid to self-employment by 72 percent.

3.2. Oaxaca-Blinder decomposition

In order to explain whether the higher earnings of employees compared with the self-employed can be explained on the basis of their attributes, we ran earnings regressions, controlling for sample selection as previously described, and ran them separately for men and women, and for employed and self-employed workers, in order to carry out the Oaxaca-Blinder decomposition[10]. The variables used in the regressions are described in detail in Table A1 in the Appendix.

Table VI below reports the estimation of the Oaxaca-Blinder decomposition of earnings differentials following the method described in Section 2.3. The first four columns in the table report the estimates and *p*-values obtained from the decomposition of the paid/self-employee differential, estimated separately for each gender. Paid employees earn more money than self-employees on an annual basis, and this differential (expressed in logs) is larger among women (0.560) than among men (0.281). According to the aggregate decomposition, the earnings gap by sector is entirely explained by selection bias into each possible outcome (non-working, self-employment, and paid employment) because the adjusted gap is negative (men) or not significantly different from zero (women). That is, if men and women were selected randomly, self-employees' earnings would be larger (men) or similar (women) than those of paid employees. Thus, the earnings gap is due neither to differences in

(10) See earnings and selection regressions in the Appendix, Tables A3 and A4.

Table VI: Oaxaca-Blinder decomposition of the employee/self-employee and male/female annual earnings gaps

	Sector earnings gap (separately for each gender)				Gender earnings gap (separately for each sector)			
	Men		Women		Self-employees		Employees	
Average annual earnings (in logs)	Estimate	P>|Z|	Estimate	P>|Z1	Estimate	P>|Z|	Estimate	P>|Z|
Employees (logs)	9.426	0.000	9.114	0.000	–	–	–	–
(€)	12.412		9.079					
Self-employees (logs)	9.145	0.000	8.553	0.000	–	–	–	–
(€)	9.360		5.184					
Men (logs)	–	–	–	–	9.145	0.000	9.426	0.000
(€)					9.360		12.412	
Women (logs)	–	–	–	–	8.553	0.000	9.114	0.000
(€)					5.184		9.079	
Unconditional gap	0.281	0.000	0.560	0.000	0.592	0.000	0.313	0.000
Adjusted gap (for selection bias)	-1.320	0.000	-0.246	0.776	1.656	0.075	0.582	0.000
Characteristics effect (explained differential)	-0.165	0.000	-0.173	0.000	0.047	0.194	0.034	0.004
Age	-0.018	0.000	-0.020	0.000	-0.004	0.426	0.015	0.000
Education	0.011	0.000	0.056	0.000	-0.019	0.108	-0.017	0.000
Experience/unemployment	-0.046	0.000	-0.072	0.000	0.007	0.706	0.036	0.000

Source: Our estimations using ECHP.

Table VI: Oaxaca-Blinder decomposition of the employee/self-employee and male/female annual earnings gaps (continuation)

	Sector earnings gap (separately for each gender)				Gender earnings gap (separately for each sector)			
	Men		Women		Self-employees		Employees	
Average annual earnings (in logs)	Estimate	P>|Z|	Estimate	P>|Z1	Estimate	P>|Z|	Estimate	P>|Z|
Industry	0.020	0.000	0.020	0.037	0.075	0.004	0.050	0.000
Occupation	-0.132	0.000	-0.164	0.000	-0.016	0.113	-0.045	0.000
Region of residence	0.002	0.323	0.009	0.002	0.003	0.007	-0.005	0.001
Wave	-0.001	0.010	-0.002	0.127	-0.001	0.731	-0.001	0.002
Coefficients effect (unexplained differential)	-1.156	0.001	-0.073	0.933	1.609	0.084	0.548	0.000
Age	0.034	0.084	-0.036	0.170	-0.067	0.032	-0.015	0.001
Education	0.033	0.049	-0.012	0.690	-0.005	0.869	-0.008	0.029
Experience/unemployment	-0.041	0.402	-0.030	0.627	-0.013	0.860	-0.026	0.006
Industry	0.001	0.954	-0.024	0.723	-0.040	0.566	0.010	0.251
Occupation	0.068	0.001	0.237	0.001	0.117	0.123	0.011	0.008
Region of residence	0.083	0.008	0.079	0.114	-0.002	0.968	0.003	0.690
Wave	-0.013	0.055	-0.041	0.001	-0.032	0.048	-0.004	0.100
Intercept	-1.321	0.001	-0.247	0.782	1.651	0.089	0.577	0.000

Source: Our estimations using ECHP.

people's endowments in the two sectors once selection has been controlled for (characteristics effect) nor to the different returns to those characteristics (coefficients effect), which are negative (or not significantly different from zero).

A negative characteristics effect for both men (-0.165) and women (-0.173) means that, roughly speaking, people working as self-employees have, in fact, better characteristics than paid employees, so that their expected earnings should be higher. This is so mostly because of the older average age of self-employees, the higher proportion of managers (occupation), and the lower share of people with previous unemployment spells. These effects predominate over their lower education (especially among men) and higher concentration in low-paid industries, such as trade or hotel and restaurants (both sexes), and transportation and real estate (men)[11].

The coefficients effect was negligible for women (-0.073), indicating that they get similar returns to their characteristics in both sectors, at least at the aggregate level. Despite this, if we look at the detailed level, we observe a positive and significant impact on the earnings differential due to a higher return to occupation in paid employment, which is compensated with the small effects of lower returns in the rest of the attributes. The coefficients effect is, however, negative and large (-1,156) in the case of men, which means that male self-employees get higher returns to their characteristics than paid employees do, and this is the consequence of a large fixed effect, which points to the omission of relevant sector-specific attributes in the regression that strongly affect earnings. Also, male paid employees get higher returns, compared with self-employees, to their age, education, and occupation, as well as to their region of residence.

Finally, the last four columns in Table V report the estimates and *p*-values of the decomposition of the male/female earnings differential, estimated separately for each sector. This data shows that the gender earnings differential is larger among self-employees (0.592) than among paid employees (0.313), and that this gap is even larger once we control for selection bias (1.656 and 0.582, respectively). Only a small part of the gender gap can be attributed to different endowments by gender in the case of employees (0.034), and nothing resulted from these differences in the case of the self-employed (0.047, but with very little significance). In the case of self-employed, only the distribution of population by industry explains a significant, even if small, part of the earnings differential while, in the case of employees, experience and age also make some limited contribution, as do education and occupation, but with a negative effect in these two cases. The larger earnings of men compared with women can be mostly attributed to a fixed effect (intercept) in both sectors.

4. Conclusions

In this paper, we have analyzed the factors that influence transitions into self-employment in Spain and the extent to which the different individual characteristics of the two groups explain why self-employees earn much less than employees.

(11) See the descriptive table of characteristics in the Appendix, Table A2.

We showed that the probability of an individual entering into self-employment depends largely on their work status. Specifically, we found that unemployed people are the most likely to become self-employed while employees are the least likely to do so. This evidence supports the view that self-employment is the only alternative for some people to become employed due to constraints in the labor market.

We also showed that some of the factors explaining the transition into self-employment differ according to previous status in the labor market. For example, while those previously out of the labor market are more likely to transition into self-employment if they have higher education, the opposite is true for employees. We interpreted this negative effect of education among employees in terms of the relationship between education and labor market opportunities. In this sense, less education is associated with greater difficulties, especially among the young, in getting permanent contracts or working in the public sector. Indeed, once we control for job characteristics, both age and education effects vanish.

In line with previous literature, we found that women are less likely to enter into self-employment regardless of their starting position, while being married increases the transition probability, although only for men out of employment. Further, wealth constraints appear to be more effective for employees than for non-workers. Lastly, the presence of children in the household and unobserved heterogeneity appear to have no effect on transitions in any case.

In short, our results for non-workers point to college-educated married men as those who are most likely to enter into self-employment in order to exit unemployment. Among workers, transition probability is lower for those with a higher opportunity cost, namely, employees in the public sector and in large firms.

Furthermore, using an Oaxaca-Blinder approach, we further rejected the hypothesis that the large earnings differential between employees and self-employees was a natural consequence of the better endowments of employees once selection was controlled for. We showed that this earnings gap is entirely the consequence of a selection bias into each possible labor status because, in fact, the adjusted gap is either negative or negligible. That is, if men and women were selected randomly, self-employees' earnings would be larger (men) or similar (women) to those of employees. Finally, we showed that the observed gender earnings gap is larger among the self-employed and could not be explained by differences in observed characteristics.

In summary, one might conclude that self-employment in Spain is mainly used as a way of entering employment for non-workers and as a way to improve working conditions for precarious workers. It is well-known that the Spanish labor market has collapsed during the ongoing economic recession, affecting both employees and self-employed, although disproportionately hitting low-skilled workers with fixed-term contracts. It is remarkable to note, however, that according to official estimates from the Labor Force Survey (EPA, INE) the number of self-employed has rebounded from 2.81 to 2.92 million between the fourth quarter of 2011 and the third of 2012. This happened in a context of a weak economy, soaring unemployment (the number of employees decreased by about 0.6 million in the same period) and unprecedented credit constraints. This highlights the relevance that en-

couraging entrepreneurship might have in the short-run as a way out of the recession even if, in the long-run, it would be useful to analyze why high-skilled workers do not see self-employment as an attractive employment alternative.

Appendix: Description variables

a) Transition regressions

Demographic characteristics (dummies):

- Education: level of education attained is less than secondary, secondary, or college.
- Age: from 19 to 30 years old, from 31 to 45, older than 45.
- Gender: the individual is a woman.
- Marital status: the individual is married.
- Children: there are children under six years old in the household.

Economic characteristics (continuous variables):

- Wealth: a proxy of household wealth as the sum of capital income and property/rental income.
- Unemployment rate of demographic group (by gender and age group).

Employment history (dummies):

- Previous labor experience: the individual has worked before.
- Tenure in the job: the individual has been over 15 years in his job.
- Size of the firm: large firms (with 50 or more employees).
- Type of sector: the individual works in the public sector. This variable interacts with the education dummies in the estimation.
- Working time: the individual works on a part-time basis.
- Type of contract: the individual works with a fixed term contract.

Additional control variables: Regional and time dummies.

b) Earnings and selection regressions

Age, education, previous experience, and time and regional control variables defined as in the transition regressions.

Selection term: estimated using a multinomial logit.

Additional labor related variables:

- Tenure (dummies): <1 year, 1-5 years, 6-15 years, 16 or more years (earnings regressions).
- Industry: manufacturing, wholesale and retail trade, hotel and restaurants, real estate and renting, services (education, health, and other social services), other activities (construction, transport, financial, public administration, etc.).

- Occupation: unskilled workers (clerks, operators, and elementary occupations), managers and professionals, technicians and associate professionals, and other occupations (craft and related trades workers; service workers, and shop and market sales workers).

Selection regressions

- Age, education, children, and time and regional control variables: as previously defined.
- Marital status: married, single, widowed, divorced.
- Capital/property income: income from capital and/or property.
- Other family income: amount of income of all the other members of the household.

Table A1: Characteristics of samples by gender and sector (in percentage of each respective group)

		Non-workers		Self-employees		Employees	
		Men	Women	Men	Women	Men	Women
Age	Age <30	45.2	30.5	14.6	18.4	25.9	33.8
	Age 31-44	30.6	38.4	51.3	47.1	48.2	48.0
	Age 45-61	24.3	31.1	34.2	34.5	25.9	18.2
Education	Less than secondary	54.3	65.2	60.9	54.4	50.8	35.7
	Secondary education	26.9	20.5	18.9	20.6	20.7	22.6
	College	18.8	14.3	20.2	24.9	28.5	41.7
Tenure	Tenure (<1 year)			7.5	9.4	17.4	20.6
	Tenure (1-5 years)			26.3	34.0	25.1	28.5
	Tenure (6-15 years)			31.7	35.6	23.8	25.3
	Tenure (16+ years)			34.6	21.0	33.7	25.5
Experience/unemployment	Previous experience			60.2	55.7	57.5	56.8
	Unemployment spell			23.9	25.1	33.5	40.1
	Long unemployment spell			11.5	19.6	13.7	23.1
Industry	Manufacture			14.0	6.2	29.6	14.0
	Trade			28.1	43.8	11.7	13.3
	Hotel - restaurants			9.8	12.4	4.3	6.1
	Real estate			10.4	10.9	4.7	9.3
	Services			5.9	23.5	11.9	39.7
	Other industry			31.9	3.2	37.8	17.6
Occupation	Unskilled			14.4	6.3	34.8	39.3
	Managers			43.2	55.0	14.5	23.0
	Technicians			8.1	5.7	11.4	12.2
	Other occupation			34.4	32.9	39.3	25.5
All		38.0	69.7	10.5	2.7	51.5	27.4

Source: Our estimations using ECHP.

Table A2: Earnings regressions

Variables	Self-employees				Paid employees			
	Men		Women		Men		Women	
	Coeff.	P>|t|	Coeff.	P>|t|	Coeff.	P>|t|	Coeff.	P>|t|
Age 31-44	-0.093	0.315	0.351	0.019	0.087	0.000	0.121	0.000
Age 45-61	-0.034	0.724	0.474	0.012	0.194	0.000	0.131	0.000
Secondary education	0.105	0.061	0.208	0.093	0.105	0.000	0.188	0.000
College	0.363	0.000	0.143	0.295	0.111	0.000	0.313	0.000
Tenure (1-5 years)	0.515	0.000	0.295	0.067	0.336	0.000	0.477	0.000
Tenure (6-15 years)	0.510	0.000	0.289	0.090	0.398	0.000	0.552	0.000
Tenure (16+ years)	0.402	0.000	-0.089	0.649	0.486	0.000	0.616	0.000
Unemployment spell	0.141	0.035	-0.327	0.112	0.043	0.000	0.104	0.000
Long unemployment spell	0.118	0.147	0.802	0.000	0.118	0.000	0.083	0.000
Experience	0.027	0.679	-0.130	0.198	-0.011	0.325	0.017	0.245
Trade	-0.131	0.046	0.102	0.599	-0.096	0.000	-0.088	0.000
Hotel – restaurants	-0.092	0.272	0.490	0.025	-0.113	0.000	-0.022	0.410
Real estate	0.085	0.354	0.278	0.238	-0.105	0.000	-0.099	0.000
Services	-0.297	0.004	0.251	0.234	-0.146	0.000	-0.147	0.000
Other industry	0.000	0.997	0.187	0.571	0.004	0.622	0.100	0.000
Unskilled	-0.050	0.454	0.264	0.199	0.019	0.022	-0.050	0.001
Managers	0.158	0.001	0.183	0.084	0.498	0.000	0.508	0.000
Technicians	0.286	0.001	0.633	0.004	0.203	0.000	0.207	0.000
Intercept	10.919	0.000	8.227	0.000	9.183	0.000	8.584	0.000
λ (selection term)	-1.130	0.000	-0.364	0.236	-0.387	0.000	-0.010	0.790
Adjusted R^2	0.101		0.080		0.438		0.470	
No. of observations	2,899		831		14,473		8,352	

Note: Time and regional dummies have been used in all regressions.

Omitted categories: 30 years old or younger person with primary education, tenure less than one year, without previous unemployment spell or labor experience, working in the manufacture sector in "other occupation", Selection terms estimated using multinomial logits.

Source: Our estimations using ECHP.

Table A3: SELECTION EQUATIONS (MULTINOMIAL LOGITS)

Variables	Men				Women			
	Self-employed		Paid employed		Self-employed		Paid employed	
	Coeff.	P>|t|	Coeff.	P>|t|	Coeff.	P>|t|	Coeff.	P>|t|
Age 31-44	0.968	0.000	0.424	0.000	0.796	0.000	0.318	0.000
Age 45-61	1.033	0.000	-0.104	0.260	0.403	0.002	-0.190	0.009
Secondary education	0.494	0.003	0.671	0.000	-0.130	0.172	0.031	0.555
College	1.008	0.000	1.699	0.000	0.187	0.067	0.642	0.000
Separated	0.080	0.805	0.829	0.000	-1.062	0.000	-1.289	0.000
Divorced	0.118	0.771	0.825	0.000	-0.535	0.177	-0.852	0.000
Widowed	0.366	0.336	0.222	0.234	-0.896	0.054	-0.814	0.009
Single	0.342	0.105	0.265	0.000	-1.302	0.000	-1.039	0.000
Children under six	-0.091	0.489	-0.282	0.000	-0.039	0.577	0.007	0.875
Other family income	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Capital/property income	0.000	0.001	0.000	0.398	0.000	0.000	0.000	0.273
Intercept	-3.415	0.000	0.369	0.002	-1.456	0.000	1.388	0.000
Pseudo R^2	0.101				0.075			
No. of observations	30,288				27,955			

Note: Time and regional dummies have been used in all regressions.

Omitted categories: 30 years old or younger married person with primary education.

Reference Group: non-employed.

Source: Our estimations using ECHP.

REFERENCES

Aguado, R., Congregado, E. and Millán, J.M. (2002): "Entrepreneurship, financiación e innovación. La situación en la Unión Europea", *Economía Industrial*, vol. 347, pp. 125-134.

Arenius, P. and Minniti, M. (2005): "Perceptual variables and nascent entrepreneurship", *Small Business Economics*, vol. 24, pp. 233-247.

Audretsch, D.B. (2007): "Entrepreneurship capital and economic growth", *Oxford Review of Economic Policy*, vol. 23, pp. 63-78.

Blanchflower, D.G. and Meyer, B. (1994): "A longitudinal analysis of the young self-employed in Australia and the United States", *Small Business Economics*, vol. 6, pp. 1-20.

Blanchflower, D.G. and Oswald, A. J. (1998): "What makes an entrepreneur?", *Journal of Labor Economics*, vol. 16, pp. 26-60.

Blanchflower, D.G. and Oswald, A.J. (1991): "Self-employment and Mrs. Thatcher's enterprise culture", CEP Working Paper 30.

Blinder, A.S. (1973): "Wage discrimination: reduced form and structural estimates", *Journal of Human Resources*, vol. 8, pp. 436-455.

Bover, O., Arellano, M. and Bentolila, S. (1996): "Duración del desempleo, duración de las prestaciones y ciclo económico", Banco de España, Estudios Económicos 57.

Cain, G.C. (1986): "The economic analysis of labor market discrimination: A survey", in Ashenfelter, O. and Layard, R. (eds.), *Handbook of Labor Economics*, vol. 1, Amsterdam: North-Holland, pp. 693-785.

Carr, D. (1996): "Two paths to self-employment? Women's and men's self-employment in the United States, 1980", *Work and Occupations*, vol. 23, pp. 26-53.

Carrasco, R. (2002): "Modelos de elección discreta para datos de panel y modelos de duración: una revisión de la literatura", *Cuadernos Económicos de ICE*, vol. 66, pp. 21-49.

Carrasco, R. (1999): "Transitions to and from self-employment in Spain: an empirical analysis", *Oxford Bulletin of Economics and Statistics*, vol. 61, pp. 315-341.

Carrasco, R. and Ejrnaes, M. (2003): "Self-employment in Denmark and Spain: Institution, economic conditions and gender differences", Centre for Applied Microeconometrics, Working Papers No. 06, University of Copenhagen.

Congregado, E., Golpe, A., Millán, J. and Román, C. (2006): "The emergence of new entrepreneurs in Europe: Which Europeans are more likely to become job creators?", Paper presented at the IX Spanish Applied Economics Meeting.

Cox, D. R. (1972): "Regression models and life tables (with discussion)", *Journal of the Royal Statistical Society*, series B, vol. 34, pp. 187-220.

De Wit, G. and Winden, F. (1989): "An empirical analysis of self-employment in the Netherlands", *Small Business Economics*, vol. 1, pp. 263-272.

Devine, T. (1994): "Changes in wage and salary returns to skill and the recent rise in female self-employment", *American Economic Review*, vol. 84, pp. 108-113.

Dunn, T. and Holtz-Eakin, D. (2000): "Financial capital, human capital, and the transition to self-employment: Evidence from intergenerational links", *Journal of Labour Economics*, vol. 18, pp. 282-305.

Evans, D.S. (1989): "Immigrant entrepreneurship: Effects of ethnic market size and isolated labor pool", *American Sociological Review*, vol. 54, pp. 950-962.

Evans, D.S. and Jovanovic, B. (1989): "An estimated model of entrepreneurial choice under liquidity constraints", *Journal of Political Economy*, vol. 97, pp. 808-827.

Evans, D.S. and Leighton, L.S. (1989): "Some empirical aspects of entrepreneurship", *American Economic Review*, vol. 79, pp. 519-535.

Fujii, E.T. and Hawley, C.B. (1991): "Empirical aspects of self-employment", *Economic Letters*, vol. 36, pp. 323-329.

Gardeazabal, J. and Ugidos, A. (2004): "More on identification in detailed wage decompositions", *Review of Economics and Statistics*, vol. 86, pp. 1034-1036.

Georgellis, Y. and Wall, H.J. (2005): "Gender differences in self-employment", *International Review of Applied Economics*, vol. 19, pp. 321-342.

Hammarstedt, M. (2004): "Self-employment among immigrants in Sweden: An analysis of intragroup differences", *Small Business Economics*, vol. 23, pp. 115-126.

Heckman, J.J. (1974): "Shadow prices, market wages and labor supply", *Econometrica*, vol. 42, pp. 679-694.

Hughes, K.D. (2003): "Pushed or pulled? Women's entry into self-employment and small business ownership", *Gender, Work and Organization,* vol. 10, pp. 433-454.

Hundley, G. (2001): "Why women earn less than men in self-employment", *Journal of Labor Research,* vol. 22, pp. 817-829.

Jenkins, S. (1995): "Easy estimation methods for discrete-time duration models", *Oxford Bulletin of Economics and Statistics*, vol. 57, pp. 129-138.

Johansson, E. (2000): "Self-employment and liquidity constraints: Evidence from Finland", *Scandinavian Journal of Economics*, vol. 102, pp. 123-134.

Kalbfleisch, J.D. and Prentice, R.L. (1980): *The Statistical Analysis of Failure Time Data*. Wiley, New York.

Kaplan, E. and Meier, P. (1958): "Non-parametric estimation from incomplete observations", *Journal of the American Statistical Association,* vol. 53, pp. 457-481.

Kidd, M.P. (1993): "Immigrant wage differentials and the role of self-employment in Australia", *Australian Economic Papers*, vol. 32, pp. 92-115.

Le, A.T. (1999): "Empirical studies of self-employment", *Journal of Economic Surveys*, vol. 13, pp. 381-416.

Lee, L.F. (1983): "Generalized econometric models with selectivity", *Econometrica*, vol. 51, pp. 507-512.

Leoni, T. and Falk, M. (2010): "Gender and field of study as determinants of self-employment", *Small Business Economics,* vol. 34, pp. 167-185.

Meyer, B. (1990): "Unemployment insurance and unemployment spells", *Econometrica*, vol. 58, pp. 757-782.

Mincer, J. (1974): *Schooling, experience and earnings*. New York: Columbia University.

Moore, R. (1983): "Employer discrimination: evidence from self-employed workers", *Review of Economics and Statistics,* vol. 65, pp. 496-501.

Oaxaca, R. (1973): "Male-female wage differentials in urban labor markets", *International Economic Review*, vol. 14, pp. 693-709.

Rees, H. and Shah, A. (1986): "An empirical analysis of self-employment in the UK", *Journal of Applied Econometrics*, vol. 1, pp. 95-108.

Suits, D.B. (1984): "Dummy variables: mechanics vs. interpretation", *Review of Economics and Statistics*, vol. 66, pp. 177-180.

Taniguchi, H. (2002): "Determinants of women's entry into self-employment", *Social Science Quarterly*, vol. 83, pp. 875-893.

Tansel, A, (2002): "Wage earners, self-employment and gender in the informal sector in Turkey", Policy Research Report on Gender and Development, Working Paper Series No. 24, The World Bank.

Taylor, M.P. (1996): "Earnings, independence, or unemployment: why become self-employed?", *Oxford Bulletin of Economics and Statistics*, vol. 58, pp. 253-266.

Tervo, H. (2006): "Regional unemployment, self-employment and family background", *Applied Economics*, vol. 38, pp. 1055-1062.

Trost, R. and Lee, L. F. (1984): "Technical training and earnings: a polytomous model with selectivity", *Review of Economics and Statistics*, vol. 66, pp. 151-156.

Yun, M-S. (2005a): "A simple solution to the identification problem in detailed wage decompositions", *Economic Inquiry*, vol. 43, pp. 766-772.

Yun, M-S. (2005b): "Normalized equation and decomposition analysis: computation and inference", Institute for the Study of Labor (IZA) Discussion Paper 1822, Bonn.

Fecha de recepción del original: abril, 2012
Versión final: diciembre, 2012

RESUMEN

En este trabajo se analizan los determinantes de la entrada en el autoempleo en España mediante la estimación de modelos de duración en tiempo discreto. Por otra parte, y ante la evidencia de menores ingresos entre los autoempleados, se realiza una descomposición de Oxaca-Blinder con el fin de explicar este diferencial. Los datos empleados en el estudio proceden del Panel de Hogares de la Unión Europea (PHOGUE) para el periodo 1994-2001. Los resultados indican que los determinantes de la entrada en el autoempleo dependen de la situación laboral del individuo. En cuanto al diferencial de ganancias, el análisis realizado indica que éste es consecuencia de un proceso de autoselección.

Palabras clave: autoempleo, datos de panel, análisis de duración, diferencial de ganancias.

Clasificación JEL: J82, J16, L26.