

Article

Estimation of Two-States Proportional Hazard Rates Models with Unobserved Heterogeneity

Emilio Congregado ^{1,†} , David Troncoso-Ponce ^{2,*,†} , Nicola Rubino ^{3,†}  and Alejandro Morales-Kirioukhina ^{1,†} 

¹ Department of Economics, CCTH (Technological and Scientific Center of Huelva), Faculty of Experimental Sciences, Campus El Carmen, University of Huelva, 21007 Huelva, Spain; congregado@uhu.es (E.C.); alejandro.morales@dege.uhu.es (A.M.-K.)

² Department of Economic Analysis and Political Economy, University of Seville, 41018 Seville, Spain

³ Department of Economic Structure, CCTH (Technological and Scientific Center of Huelva), Faculty of Experimental Sciences, Campus El Carmen, University of Valencia, 21007 Huelva, Spain; nicola.rubino@uv.es

* Correspondence: dtroncoso@us.es

† These authors contributed equally to this work.

Abstract

This article examines two-state proportional hazard rate models with unobserved heterogeneity specific to each state, a framework that is especially relevant for labor market transitions. To make estimation feasible in large longitudinal datasets, we implement `hshaz2s`, a Stata routine that uses analytical expressions for the gradient vector and Hessian matrix of the log-likelihood function through the dual second-order moment (`d2_m1`) method. The empirical application estimates a discrete-time duration model for transitions between employment and unemployment using Spanish labor market microdata for young low-skilled workers over 2000–2019. The results show that apprenticeship contracts are associated with lower exit rates from employment than other temporary contracts, but not with faster transitions from unemployment back into employment. The estimates also reveal substantial state-specific unobserved heterogeneity, with a large latent group characterized by persistent spells in both states. Analytical second-order information also markedly reduces convergence time under richer heterogeneity structures. Overall, the article makes this class of two-state hazard models operational for applied research and provides new evidence on apprenticeship and temporary contracts in Spain.

Keywords: duration models; unobserved heterogeneity; `d2_m1` method; `hshaz2s`; gradient vector; Hessian matrix; multistate duration models; proportional hazard rates models



Academic Editor: Julien Chevallier

Received: 27 December 2024

Revised: 9 April 2026

Accepted: 16 April 2026

Published: 28 April 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Survival analysis has become a standard framework for studying time-to-event processes in economics and other applied fields. In medical research, this approach is used to analyze the timing of events such as disease progression, recovery, and recurrence, while in labor economics it is widely applied to employment duration, job retention, and re-employment processes (Lechner & Pfeiffer, 2001; Machado & van den Hout, 2018). This line of research is especially relevant when the objective is to understand labor market instability, job retention, and re-employment prospects in settings characterized by recurrent mobility between employment and unemployment. Recent work has continued to emphasize the importance of labor force transitions as a central dimension of labor market dynamics and has highlighted the substantial heterogeneity underlying these flows (Castro et al., 2024).

A substantial part of the empirical literature on labor market duration analysis has traditionally relied on one-equation models that estimate transition rates separately for each origin state. This is the case of studies that analyze exits from employment and unemployment independently using single-equation duration models (see, for example, [Alba-Ramírez et al., 2007, 2012](#); [Arranz & García-Serrano, 2014](#); [Rebollo-Sanz & García-Pérez, 2015](#)). More recent contributions, however, have increasingly favored the joint estimation of hazard rates from both employment and unemployment states, thereby allowing the researcher to model labor market dynamics in a more integrated way (see, for example, [Arranz et al., 2010](#); [Bentolila et al., 2017](#); [Felgueroso et al., 2018](#); [Kyyra et al., 2019](#)). In general, labor economics is concerned not only with the marginal effects of observed covariates on spell duration, but also with the way individuals move repeatedly across multiple states, such as employment, unemployment, and non-participation. Multi-state models are particularly useful in this context because they provide a richer framework for representing such transitions and their dependence structure ([Bentolila et al., 2017](#); [Felgueroso et al., 2018](#); [Kyyra et al., 2019](#); [van den Hout & Tan, 2019](#)). Recent evidence has also reinforced the importance of accounting for heterogeneity in unemployment duration and labor market transitions, as well as the role of temporary employment arrangements in shaping subsequent trajectories ([Ahn, 2023](#); [Carrasco et al., 2024](#); [Eberlein et al., 2024](#)).

An important advantage of this framework is that it allows unobserved heterogeneity to be modeled in a way that is specific to each state. In labor market applications, individuals may differ in persistent but unobserved traits affecting both job stability and re-employment prospects, such as motivation, search intensity, adaptability, or match quality. Ignoring these factors may bias the estimated transition rates and distort the interpretation of duration dependence. For this reason, the joint estimation of transition hazards with state-specific unobserved heterogeneity is of clear interest in applied work. More generally, multi-state duration models have been shown to provide a more nuanced representation of transition processes by accommodating richer latent structures and more flexible dependence patterns ([Bieszk-Stolorz, 2021](#); [van den Hout & Tan, 2019](#)). Recent work has also shown that unobserved heterogeneity remains central for understanding unemployment duration and labor market persistence ([Ahn, 2023](#)). At the same time, the growing literature on temporary employment and youth labor market insertion suggests that early contractual arrangements may shape subsequent employment trajectories in complex ways ([Eberlein et al., 2024](#); [Vaquero García et al., 2024](#)).

The usefulness of multi-state models is not confined to labor economics. In medical and biostatistical applications, they have also been employed to study transitions across health states, including disease progression, recovery, and interval-censored outcomes, showing the value of flexible specifications that account for latent heterogeneity and complex event histories ([Machado & van den Hout, 2018](#)). These contributions underscore the broader methodological relevance of multi-state duration analysis, but they also make clear that richer models typically require more demanding estimation procedures and greater computational effort.

Despite this progress, empirical implementation still tends to be tailored to specific applications, and readily deployable tools for jointly estimating employment and unemployment hazards with state-specific unobserved heterogeneity remain limited in standard applied software, especially when working with large monthly-expanded administrative datasets. Against this background, the contribution of this article is twofold. First, it operationalizes the estimation of two-state proportional hazard rate models with unobserved heterogeneity specific to each state, allowing for a flexible bivariate discrete-mixture structure in the spirit of [Heckman and Singer \(1984\)](#). Second, it provides an implementation in Stata through the `hshaz2s` command, which incorporates analytical expressions for the

gradient vector and the Hessian matrix of the log-likelihood function using the dual second-order moment (d2 ml) method (Gould et al., 2010). This improves numerical efficiency and makes the estimation of these models feasible in large-scale longitudinal datasets.

The empirical relevance of the proposed framework is illustrated through an application to transitions between employment and unemployment in a sample of young low-skilled workers in the Spanish labor market. This setting is particularly appropriate because it involves repeated labor market mobility and strong heterogeneity in employment stability and job-finding prospects. The application focuses on the differential transition patterns associated with apprenticeship contracts relative to other temporary contracts, while jointly modeling exits from employment and unemployment. This focus is also timely in light of recent comparative evidence on training contracts and youth labor market integration in Spain and other European countries (Cueto & López, 2019; Jansen & Troncoso-Ponce, 2018; Vaquero García et al., 2024).

The remainder of the article is organized as follows. Section 2 describes the longitudinal database and the econometric specification. Section 3 reports the estimation results, including the evidence on unobserved heterogeneity and the computational gains from the use of analytical derivatives. Section 4 discusses the main empirical findings. Section 5 concludes.

2. Materials and Methods

2.1. Data: The Continuous Sample of Working Histories

The empirical application is based on the Continuous Sample of Working Histories (CSWH), a longitudinal administrative database for the Spanish labor market. Following Troncoso-Ponce (2017) and Troncoso-Ponce (2018), we use this source to construct a sample of 132,262 young low-skilled workers observed over the period 2000–2019. The CSWH contains complete labor market histories for more than one million individuals and represents a 4% non-stratified random draw from the population linked to the Spanish Social Security Administration. It includes both wage earners and recipients of Social Security benefits, such as unemployment benefits, disability benefits, survivor pensions, and maternity leave (see, for example, Arranz & García-Serrano, 2011; Lafuente, 2020; Lapuerta, 2010).

The database provides the start and end dates of employment and unemployment episodes throughout the observed labor history of each individual, together with personal and job-related information. This structure makes it possible to reconstruct monthly labor market spells and to estimate transition rates between employment and unemployment in a longitudinal framework. Additional details on the variables available in the CSWH can be found in García-Pérez (2008) and Arranz et al. (2013). This type of administrative source has also continued to prove useful in recent research on labor market transitions in Spain based on discrete-time duration frameworks and rich employment histories (Carrasco et al., 2024).

The estimation sample consists of young workers with low educational attainment and qualifications in the Spanish labor market between 2000 and 2019. Their average age is 21.9 years and 36.7% are women. On average, each worker experiences 6.4 employment episodes and 6.7 unemployment episodes, with mean durations of 10.3 and 11.7 months, respectively. Approximately 25% of these episodes last no more than 2 months in unemployment and 3 months in employment, whereas 5% extend to at least 37 and 35 months, respectively, indicating both substantial turnover and persistent non-employment spells. After expanding the spell data to the monthly level and redefining the time-varying information accordingly, the estimation sample contains 8,001,341 person-month observations, of which 4,485,930 correspond to unemployment and 3,515,411 to employment. Appendix A.2 reports the main descriptive statistics.

2.2. Econometric Model

The empirical framework jointly estimates monthly transition rates from two mutually exclusive labor market states: employment and unemployment. Individuals are followed over the observation period and, at each month, can be observed in one of these two states (see Allison, 1982; Jenkins, 1995; Lancaster, 1992). The model is specified in discrete time and allows the transition process to differ across origin states, both in terms of observed covariates and unobserved heterogeneity.

Consider an individual who begins an (un)employment episode at time $T = 1$, where time is measured in monthly intervals. The individual is then observed month by month until either a transition occurs from the current state to the destination state modeled for that equation or the observation period ends, in which case the spell is right-censored.

The hazard rate out of (un)employment is modeled as follows:

$$h^s(t|x^s, \eta^s) = 1 - \exp(-\exp(\lambda^s(t) + x^s\beta^s + \eta^s)) \quad (1)$$

where $s = \{u, e\}$, with u denoting unemployment and e denoting employment. The hazard rate at month $T = t$ depends on duration dependence, captured by $\lambda^s(t)$, on a vector of observed covariates x^s , whose effects are summarized by β^s , and on a state-specific unobserved component η^s .

In both equations, duration dependence is modeled flexibly through sets of duration dummies. In the employment equation, the specification is further allowed to differ by contract type. More specifically, the employment equation is specified so as to allow the duration profile of apprenticeship contracts to differ from that of other temporary contracts, in line with the evidence suggested by non-parametric Kaplan–Meier estimates. This choice is intended to capture the possibility that apprenticeship contracts follow a distinct employment-duration pattern relative to the rest of temporary contracts. The vector x^s includes both common and state-specific regressors. In the employment equation, the covariates capture individual characteristics, type of labor contract, sector of activity, regional conditions, and calendar-time effects. In the unemployment equation, the covariates include individual characteristics, characteristics of the previous job spell, regional labor market conditions, and calendar-time effects. In both cases, the specification allows for time-varying covariates whenever the underlying information changes over the spell. This flexible formulation is intended to isolate duration dependence from compositional differences across workers and labor market conditions.

A key feature of the proposed framework is that, unlike the earlier `hshaz2` and `hshaz` commands, which are restricted to the estimation of hazard rates from a single state, it allows the joint estimation of transition rates from both states within a unified specification and, by applying the methodology proposed by Heckman and Singer (1984), accommodates the identification of unobserved heterogeneity as a non-parametric bivariate discrete mixture of state-specific latent effects η^s . The contribution to the total likelihood function for an individual i with unobserved heterogeneity captured by the vector (η_i^u, η_i^e) is given by:

$$\mathcal{L}_i(\eta_i^u, \eta_i^e) = \left(\prod_{t=1}^{T_i^u} h^u(t|\lambda^u(t), x_{it}^u, \eta_i^u)^{y_{it}^u} S^u(t-1|\lambda^u(t-1), x_{it-1}^u, \eta_i^u)^{(1-y_{it}^u)} \right)^{I(u_{it}=1)} \left(\prod_{t=1}^{T_i^e} h^e(t|\lambda^e(t), x_{it}^e, \eta_i^e)^{y_{it}^e} S^e(t-1|\lambda^e(t-1), x_{it-1}^e, \eta_i^e)^{(1-y_{it}^e)} \right)^{I(e_{it}=1)} \quad (2)$$

where $h^u(t|\lambda^u(t), x_{it}^u, \eta_n^u)$ and $h^e(t|\lambda^e(t), x_{it}^e, \eta_m^e)$ denote the hazard rates out of unemployment and employment at month $T = t$, respectively. Likewise, $S^u(t-1|\lambda^u(t-1), x_{it-1}^u, \eta_n^u)$ and $S^e(t-1|\lambda^e(t-1), x_{it-1}^e, \eta_m^e)$ denote the survivor functions at month $T = t-1$ in the unemployment and employment states. As shown in Equation (2), both the hazard rates and the survivor functions depend on state-specific observed and unobserved covariates, some of which may vary over time.

Unobserved heterogeneity is introduced through a discrete-mixture specification that is allowed to differ across the unemployment and employment equations. The latent effects are represented by a finite number of support points for each state, and their joint distribution is modeled as a bivariate discrete mixture. Identification is achieved from the longitudinal structure of the data and the joint estimation of transitions from both states. Since individuals contribute repeated employment and unemployment episodes over time, the model exploits within-individual variation in spell durations, censoring patterns, and transition histories to recover the support points and their associated probabilities.

The number of support points used to approximate the distribution of unobserved heterogeneity must balance flexibility and parsimony. In practice, models with 2, 4, or 9 mass points can be estimated, corresponding to increasingly rich discrete approximations of the latent distribution. Model choice should be guided by a combination of statistical fit and practical interpretability. In particular, improvements in the log-likelihood and information criteria, together with the stability of the estimated support points and their probabilities, provide a natural basis for selection (see, for example, Gaure et al., 2007; Nicoletti & Rondinelli, 2010). When additional mass points generate only marginal gains in fit, very small estimated probabilities, or substantially higher computational costs without altering the substantive conclusions, the more parsimonious specification is preferable. In the empirical application below, the four-point specification offers a useful compromise between flexibility, interpretability, and computational feasibility.

The total likelihood function for the empirical specification used in this article is:

$$\mathcal{L} = \prod_{i=1}^N \sum_{n=1}^2 \sum_{m=1}^2 \pi_{nm} \mathcal{L}_i(\eta_n^u, \eta_m^e) \quad (3)$$

with $\pi_{nm} = \frac{e^{p_{nm}}}{1 + e^{p_{12}} + e^{p_{21}} + e^{p_{22}}}$, where $nm = \{12, 21, 22\}$. Equation (3) corresponds to the

four-point specification used in the baseline empirical exercise. More generally, the same framework can be extended to alternative numbers of support points by enlarging the discrete support of the latent effects in each state and reparameterizing the associated probability masses accordingly.

The model parameters are estimated by maximum likelihood, using analytical expressions for the gradient vector and the Hessian matrix of the log-likelihood function. This makes it possible to implement the `d2 ml` method for the maximization of $\ln \mathcal{L}$ and improves numerical efficiency in the estimation of the two-state duration model.

In the employment equation, the covariates include: (i) personal characteristics (sex, age, age squared, education, and nationality); (ii) business-cycle conditions, measured by the quarterly regional unemployment rate and its interactions with the elapsed duration of the spell; (iii) regional fixed effects; (iv) a fully non-parametric baseline hazard defined through employment-duration dummies; and (v) characteristics of the current job, including the type of contract and the sector of activity. In the unemployment equation, the specification includes: (i) the same set of personal characteristics; (ii) the quarterly regional unemployment rate; (iii) regional fixed effects; (iv) a fully non-parametric baseline hazard defined through unemployment-duration dummies; and (v) characteristics of the

previous job spell, including the type of contract previously held and the sector of the last job. Descriptive statistics for both sets of regressors are reported in Appendix A.2.

3. Results

This section reports the empirical results obtained from the estimation of the two-state discrete-time duration model with unobserved heterogeneity using a sample of young low-skilled workers in the Spanish labor market over the period 2000–2019. Earlier versions of this sample, covering 2000–2014, were analyzed in Troncoso-Ponce (2017) and Troncoso-Ponce (2018) using the `hshaz2` and `hsmlogit` commands, respectively. The present specification extends that previous work by jointly estimating transition rates out of both employment and unemployment within a unified framework.

The application centers on whether apprenticeship contracts are associated with different transition patterns from those observed for other temporary contracts. Apprenticeship contracts are temporary contracts targeted at young workers aged 16–30 and at low-qualified workers without a university degree or completed vocational training. Their institutional design includes an officially approved training scheme intended to provide technical knowledge and work experience for a qualified occupation. These contracts last between six months and three years and are supported by public subsidies that help employers finance the training component.

To study these differences, the analysis distinguishes between two contract categories, apprenticeship contracts and the remaining temporary contracts, and tracks workers not only during the employment spell in which the contract is observed, but also during subsequent employment and unemployment episodes over the observation window. In this way, the analysis compares transition rates out of employment and out of unemployment according to whether the current or previous employment relationship corresponds to an apprenticeship contract or to another temporary contract. Although the terminology of treatment and control groups may be useful descriptively, the purpose of the exercise is not to identify causal effects, but rather to document differences in transition patterns within the joint hazard framework described in Section 2.2.

In both the employment and unemployment equations, duration dependence is modeled fully non-parametrically through sets of duration dummies, and the specification also controls for personal characteristics, job-related characteristics, regional conditions, and the economic cycle. The full set of covariates is reported in Appendix A.2, while Appendix A.1 presents the estimated coefficients.

Figure 1 reports the predicted mean hazard rates for exits from employment and unemployment. The left panel shows the employment hazard rates for workers holding apprenticeship contracts and for workers holding other temporary contracts. In both groups, the probability of leaving employment declines with tenure, although the profile is not identical across contract types. At almost all observed durations, workers employed under apprenticeship contracts display lower exit rates from employment than workers employed under other temporary contracts. This pattern is also consistent with earlier evidence on the prevalence of very short temporary contracts in the Spanish labor market (Felgueroso et al., 2018). The gap is especially visible during the first months of the employment spell, when separation risks are highest overall.

These differences are consistent with the coefficient estimates reported in Appendix A.1. In the employment equation, the coefficient on Apprenticeship contract (=1) is negative and highly significant, indicating that apprenticeship contracts are associated with a lower probability of transition out of employment relative to the reference category of temporary contracts. The estimated duration effects also show that exit rates from employment are highest at short durations and then tend to decline, although not monotonically, which is

compatible with a labor market characterized by temporary contracts and contract-renewal margins at specific durations.

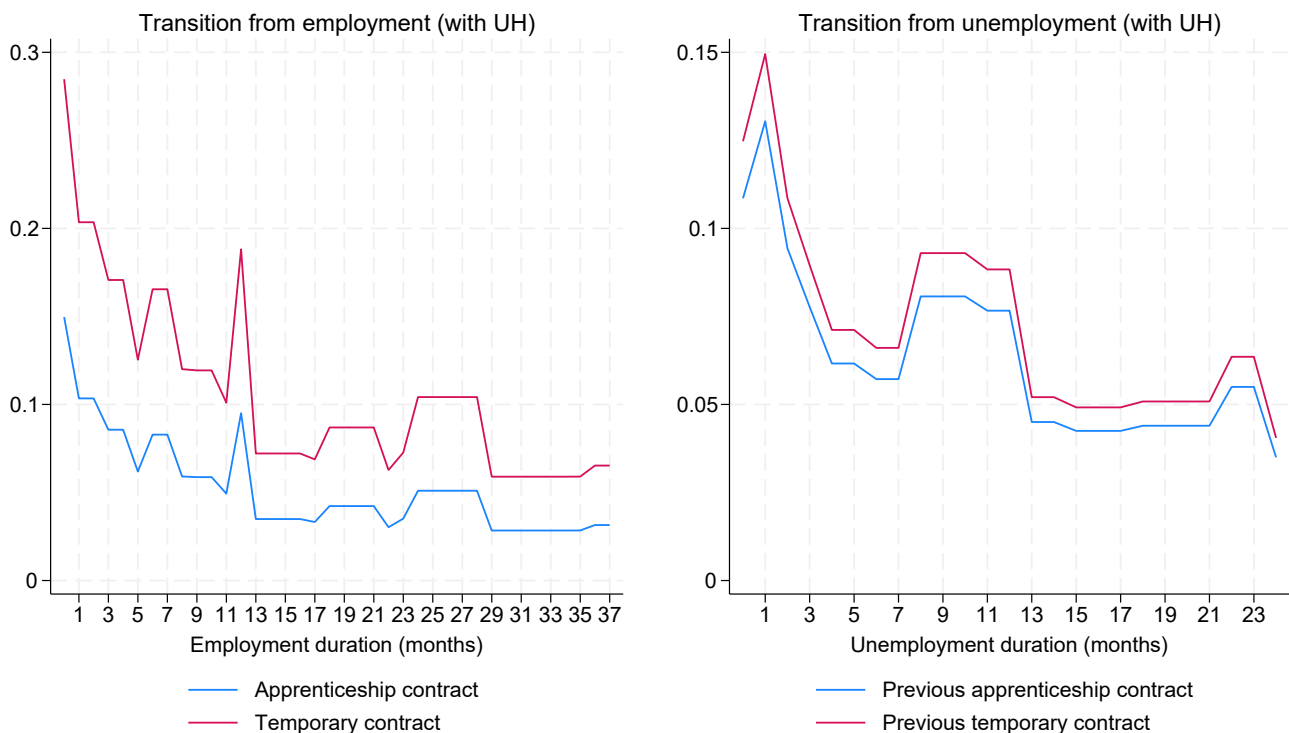


Figure 1. Employment and unemployment mean predicted hazard rates (with UH).

By contrast, the right panel of Figure 1 shows that, once workers move into unemployment, those whose previous contract was an apprenticeship contract face lower re-employment hazards than unemployed workers coming from other temporary contracts. This pattern is again consistent with the coefficient estimates: in the unemployment equation, the coefficient on Previous job: Apprenticeship contract (=1) is negative and highly significant, indicating a lower probability of exit from unemployment for workers with a previous apprenticeship episode. In both groups, the unemployment hazard profile displays a broadly declining pattern with elapsed unemployment duration, consistent with negative duration dependence in job finding. A possible interpretation, already suggested in the literature on apprenticeship contracts in Spain, is that workers with a previous apprenticeship episode may become more selective in their job search once the training spell ends (Cueto & López, 2019; Jansen & Troncoso-Ponce, 2018).

Taken together, the estimated hazard profiles reveal a clear asymmetry. Apprenticeship contracts are associated with lower transition rates out of employment, but not with faster transitions out of unemployment once the employment relationship ends. This result is relevant for the analysis of apprenticeship arrangements in the Spanish labor market, since it suggests that employment stability during the spell and subsequent re-employment prospects need not move in the same direction.

3.1. Unobserved Heterogeneity

The estimated model allows unobserved heterogeneity to differ across the unemployment and employment equations through a bivariate discrete-mixture specification. In contrast with the earlier single-state formulations developed in `hshaz2` and `hsmlogit`, where the latent structure was represented through a univariate discrete mixture, the present framework jointly identifies combinations of latent traits that are specific to the two

modeled states (Troncoso-Ponce, 2017, 2018). In contrast with earlier single-state formulations, where the latent structure is represented by a univariate discrete mixture, the present framework jointly identifies combinations of latent traits that are specific to each of the two modeled states. In the empirical specification used here, the distribution of unobserved heterogeneity is approximated by four points of support, corresponding to the combinations (Type I unemployed, Type I employed), (Type I unemployed, Type II employed), (Type II unemployed, Type I employed), and (Type II unemployed, Type II employed).

Table 1 reports the estimated support points and their associated probabilities. The results indicate substantial heterogeneity in latent transition propensities across workers. The first group, representing 9.89% of the sample, is characterized by comparatively higher exit rates from both unemployment and employment. The second group, accounting for 8.78% of the sample, combines relatively fast exits from unemployment with lower exit rates from employment. The third group, representing 17.1% of the sample, displays the opposite pattern: lower exit rates from unemployment and higher exit rates from employment. Finally, the fourth group, which accounts for 64.23% of the sample, is characterized by relatively persistent spells in both states, with lower transition rates out of both employment and unemployment.

These estimates show that the latent structure of labor market transitions is not well summarized by a single unobserved factor. Rather, the results support a specification in which unobserved heterogeneity is state-specific and jointly distributed across employment and unemployment transitions. Figure 2 complements Table 1 by displaying the predicted hazard profiles associated with the different latent types.



Figure 2. Employment and unemployment mean predicted hazard rates with UH (by Type of workers).

Table 1. Unobserved heterogeneity estimates.

	Coef.	Std. Error
η_1^u	−0.501	0.0325
η_2^u	−1.479 ^a	0.00463
η_1^e	−1.410	0.0332
η_2^e	−2.435 ^b	0.00343
$Pr(\eta_1^u, \eta_1^e)$	9.89%	0.00165
$Pr(\eta_1^u, \eta_2^e)$	8.78%	0.00234
$Pr(\eta_2^u, \eta_1^e)$	17.1%	0.00302
$Pr(\eta_2^u, \eta_2^e)$	64.23%	0.00327

^a (= −0.501 − 0.978) ^b (= −1.410 − 1.025).

3.2. Interpretation of Unobserved Heterogeneity Estimates

The parameterization of the latent structure follows the same logic as in `hshaz`, `hshaz2`, and `hsmlogit` (Troncoso-Ponce, 2017, 2018). In the `hshaz2s` output, the estimates of η_2^u and η_2^e are reported as deviations from the baseline latent components of the unemployment and employment equations, respectively. Thus, in the unemployment equation, the coefficient −0.978 reported in Table 1 corresponds to the differential latent effect of Type II unemployed relative to Type I unemployed, whose baseline latent component is captured by the constant term of the unemployment equation, −0.501. Analogously, in the employment equation, the coefficient −1.025 measures the differential latent effect of Type II employed relative to Type I employed, whose baseline latent component is given by the constant term of the employment equation, −1.410.

Under this parameterization, the support points reported in Table 1 are obtained by combining the baseline latent component and the corresponding differential effect in each equation. Table 2 reports the general structure of the bivariate discrete-mixture distribution for the alternative two-, four-, and nine-point specifications that can be estimated within the proposed framework.

Table 2. Unobserved heterogeneity distribution with two/four/nine points of support.

UH distribution with two points of support			
$\eta^u \setminus \eta^e$	η_1^e	η_2^e	
η_1^u	$\pi_{11} = Pr(\eta_1^u, \eta_1^e)$	—	
η_2^u	—	$\pi_{22} = Pr(\eta_2^u, \eta_2^e)$	
Likelihood function	$\mathcal{L} = \prod_{i=1}^N \{ \pi_{11} \mathcal{L}_i(\eta_1^u, \eta_1^e) + \pi_{22} \mathcal{L}_i(\eta_2^u, \eta_2^e) \}$		
Mass-point probabilities	$\pi_{11} = 1 - \pi_{22}$ $\pi_{22} = \frac{e^{p_{22}}}{1 + e^{p_{22}}}$		
UH distribution with four points of support			
$\eta^u \setminus \eta^e$	η_1^e	η_2^e	
η_1^u	$\pi_{11} = Pr(\eta_1^u, \eta_1^e)$	$\pi_{12} = Pr(\eta_1^u, \eta_2^e)$	
η_2^u	$\pi_{21} = Pr(\eta_2^u, \eta_1^e)$	$\pi_{22} = Pr(\eta_2^u, \eta_2^e)$	
Likelihood function	$\mathcal{L} = \prod_{i=1}^N \sum_{n=1}^2 \sum_{m=1}^2 \pi_{nm} \mathcal{L}_i(\eta_n^u, \eta_m^e)$		
Mass-point probabilities	$\pi_{11} = 1 - \pi_{12} - \pi_{21} - \pi_{22}$ $\pi_{nm} = \frac{e^{p_{nm}}}{1 + e^{p_{12}} + e^{p_{21}} + e^{p_{22}}}$		
UH distribution with nine points of support			
$\eta^u \setminus \eta^e$	η_1^e	η_2^e	η_3^e
η_1^u	$\pi_{11} = Pr(\eta_1^u, \eta_1^e)$	$\pi_{12} = Pr(\eta_1^u, \eta_2^e)$	$\pi_{13} = Pr(\eta_1^u, \eta_3^e)$
η_2^u	$\pi_{21} = Pr(\eta_2^u, \eta_1^e)$	$\pi_{22} = Pr(\eta_2^u, \eta_2^e)$	$\pi_{23} = Pr(\eta_2^u, \eta_3^e)$
η_3^u	$\pi_{31} = Pr(\eta_3^u, \eta_1^e)$	$\pi_{32} = Pr(\eta_3^u, \eta_2^e)$	$\pi_{33} = Pr(\eta_3^u, \eta_3^e)$
Likelihood function	$\mathcal{L} = \prod_{i=1}^N \sum_{n=1}^3 \sum_{m=1}^3 \pi_{nm} \mathcal{L}_i(\eta_n^u, \eta_m^e)$		
Mass-point probabilities	$\pi_{11} = 1 - \pi_{12} - \pi_{13} - \pi_{21} - \pi_{22} - \pi_{23} - \pi_{31} - \pi_{32} - \pi_{33}$ $\pi_{nm} = \frac{e^{p_{nm}}}{1 + e^{p_{12}} + e^{p_{13}} + e^{p_{21}} + e^{p_{22}} + e^{p_{23}} + e^{p_{31}} + e^{p_{32}} + e^{p_{33}}} , \quad (n, m) \neq (1, 1)$		

3.3. Computation Time Under d1 m1 and d2 m1

As discussed in Section 1, the estimation strategy incorporates analytical expressions for both the gradient vector and the Hessian matrix, which allows the use of the d2 m1 method (Gould et al., 2010). This is especially relevant when estimating two-state duration models on large monthly-expanded longitudinal datasets and extends the computational logic already exploited in earlier single-state implementations (Troncoso-Ponce, 2017).

Table 3 reports the convergence times obtained under alternative specifications with two, four, and nine mass points, comparing the d1 m1 and d2 m1 methods. All models include the same set of eighty-one covariates (thirty-seven in the unemployment equation and forty-four in the employment equation), covering duration dependence, individual characteristics, contract type, regional fixed effects, and regional macroeconomic conditions.

The results show that the use of d2 m1 substantially reduces convergence time in all cases. For the model with two mass points, the estimation time falls from 1 h and 20 min under d1 m1 to just over 5 min under d2 m1. For the model with four mass points, the required time declines from 2 h and 50 min to less than 15 min. For the most complex specification, with nine mass points, the contrast is even stronger: estimation time falls from almost 13 h to 1 h and 20 min. A second relevant result is that the convergence time under d2 m1 increases much less sharply with model complexity than under d1 m1. Whereas the difference between the simplest and the most demanding specification is more than 11 h under d1 m1, it is only about 1 h and 15 min under d2 m1. These results confirm that analytical second-order information is particularly valuable when the latent structure becomes richer.

Table 3. Duration needed to estimate two-states proportional hazard rate models (Sample size: 8,001,341 observations).

Num. of Mass-Points	Time (hh:mm:ss)		
	d1 Method	d2 Method	Diff. = d1 – d2
Two mass-points	1:20:22	0:05:13	1:15:09
Four mass-points	2:50:13	0:14:55	2:35:18
Nine mass-points	12:55:07	1:20:13	11:34:54

4. Discussion

The empirical results show that jointly modeling transitions out of employment and unemployment provides a more informative picture of labor market dynamics than estimating each process separately. In the present application, the estimated hazard profiles indicate that workers holding apprenticeship contracts face systematically lower transition rates out of employment than workers under other temporary contracts. This result is also consistent with the coefficient estimates reported in Appendix A.1: the coefficient on Apprenticeship contract (=1) in the employment equation is negative and highly significant (−0.751), indicating a substantially lower probability of exit from employment relative to the reference category of temporary contracts. In substantive terms, this suggests that apprenticeship contracts are associated with lower separation rates during employment spells.

This lower exit risk from employment is particularly relevant in light of the duration profile estimated for employment spells. The coefficients on the employment-duration dummies are generally positive at short durations and decline over time, although not in a strictly monotonic way. Exit rates are especially high at the beginning of the employment spell, with large and significant coefficients in month 1 and in months 2 to 5, and remain elevated around some contract-relevant thresholds, such as month 13 and months 25 to

29. This pattern is consistent with an institutional setting in which temporary employment relationships are often shaped by short expected durations, renewal margins, and fixed-term contractual thresholds, a feature that has been widely documented for the Spanish labor market (Felgueroso et al., 2018). Against this background, the lower separation rates associated with apprenticeship contracts suggest that this contractual arrangement provides a more stable employment relationship than standard temporary contracts, at least while the match remains active.

By contrast, the unemployment equation points in the opposite direction regarding post-unemployment transitions. The coefficient on Previous job: Apprenticeship contract (=1) is negative and highly significant (-0.158), which indicates that unemployed workers coming from an apprenticeship contract display a lower hazard of re-employment than those coming from other temporary contracts. This result is in line with the predicted hazard rates reported in Figure 1, where the unemployed with a previous apprenticeship episode exhibit lower exit rates from unemployment. Therefore, the empirical evidence suggests that apprenticeship contracts are associated with greater stability while workers remain employed, but not with faster re-entry into employment once the apprenticeship spell ends.

This asymmetry is important from an economic point of view. It suggests that employment retention and re-employment prospects should not be interpreted as equivalent dimensions of labor market performance. Apprenticeship contracts may strengthen the initial match through the training component of the contract, better match quality, or stronger retention incentives during the employment relationship. However, once the spell ends and the worker returns to unemployment, the acquired training does not necessarily translate into a higher probability of immediate re-employment in the external labor market. As suggested in previous studies, one possible explanation is that workers with prior apprenticeship experience become more selective in their search process, either because they expect better job matches or because their reservation wages increase after the training period (Cueto & López, 2019; Jansen & Troncoso-Ponce, 2018). More generally, the results indicate that apprenticeship contracts may improve stability within employment without necessarily reducing subsequent unemployment duration.

The coefficient estimates also confirm that the transition process is shaped by a broad set of observed characteristics. In the unemployment equation, being female reduces the probability of exit from unemployment, while age initially increases it and then does so at a decreasing rate, as indicated by the positive coefficient on age and the negative coefficient on age squared. Non-Spanish nationality is associated with a higher unemployment exit rate, whereas higher regional unemployment significantly lowers the probability of re-employment. In the employment equation, women display a higher hazard of exit from employment, lower educational attainment is associated with higher separation risk, and non-Spanish nationality also increases the probability of job loss. Regional labor market conditions are again important, with the quarterly regional unemployment rate increasing the hazard of exit from employment, although its interaction with spell duration reveals that this effect is not constant over time. These patterns reinforce the importance of modeling both equations jointly while allowing the impact of observed covariates to differ by origin state.

The duration coefficients in the unemployment equation reveal a pattern broadly consistent with negative duration dependence in job finding. The estimated hazards are highest in the early months of unemployment, particularly in months 1 and 2, and then tend to decline as unemployment duration lengthens, despite some local non-monotonicities. This suggests that the probability of re-employment is concentrated relatively early in the unemployment spell and becomes lower as the spell persists. In the employment equation,

the hazard of leaving employment is also strongest at short durations, but the profile is more irregular than in unemployment. This is plausible in a labor market with substantial contractual segmentation, where exits may cluster around administratively or contractually meaningful durations. For this reason, the decision to model duration dependence fully non-parametrically through duration dummies appears especially appropriate, since it allows the data to capture these irregularities without imposing an overly restrictive functional form.

A central contribution of the article concerns the estimated distribution of unobserved heterogeneity. Building on the latent-heterogeneity logic developed in earlier single-state implementations (Troncoso-Ponce, 2017, 2018), the bivariate discrete-mixture specification identifies four latent groups that differ in their propensity to leave employment and unemployment. The bivariate discrete-mixture specification identifies four latent groups that differ in their propensity to leave employment and unemployment. The estimates reported in Table 1 show that the largest group, representing 64.23% of the sample, is characterized by lower exit rates from both unemployment and employment, that is, by relatively persistent spells in both states. At the opposite end, a much smaller group, representing 9.89% of the sample, is characterized by comparatively higher transition rates out of both states. The remaining groups display asymmetric patterns, with lower exit rates in one state but higher transition rates in the other. These results indicate that unobserved heterogeneity is not simply a residual nuisance term, but a substantive dimension of labor market dynamics. Workers differ not only in observed demographic and job-related characteristics, but also in latent traits that shape their job stability and re-employment prospects in distinct ways across states.

This finding is important for the interpretation of the estimated duration effects and contract effects. If unobserved heterogeneity were omitted, part of the estimated duration dependence could reflect compositional sorting across workers rather than genuine state dependence. Likewise, the effects attributed to apprenticeship and temporary contracts could be confounded by persistent latent differences in employability, search behavior, or job-match quality. By allowing unobserved heterogeneity to be state-specific and jointly distributed across the two equations, the model provides a more credible representation of the underlying transition process and a richer interpretation of labor market segmentation among young low-skilled workers, in line with the motivation for flexible discrete-mixture approaches in duration analysis (Heckman & Singer, 1984).

The discussion of the results also clarifies the role of the computational contribution of the paper. The value of the analytical gradient vector and Hessian matrix is not merely technical, nor does it constitute an end in itself. Its importance lies in making it feasible to estimate a substantively meaningful two-state duration model with flexible unobserved heterogeneity in a very large monthly-expanded longitudinal sample, while exploiting the computational gains associated with second-order maximum-likelihood optimization (Gould et al., 2010). The strong reduction in convergence time documented in Table 3, especially as the number of mass points increases, is therefore relevant because it enables the practical estimation of a specification that would otherwise be much more difficult to implement in applied work.

Overall, the empirical evidence points to a nuanced interpretation of apprenticeship contracts in the Spanish labor market. On the one hand, they appear to reduce the risk of separation during employment spells, suggesting a stabilizing effect on early labor market trajectories. On the other hand, once workers become unemployed, prior apprenticeship experience is not associated with faster transitions back into employment. At the same time, the estimated bivariate discrete-mixture structure shows that labor market trajectories are shaped by substantial state-specific unobserved heterogeneity, over and above the rich set

of observed covariates included in the model. Taken together, these findings support the usefulness of a joint two-state framework for analyzing employment and unemployment transitions and respond directly to the need for a more substantive interpretation of the empirical results.

5. Conclusions

This article examines the estimation of two-state proportional hazard rate models with unobserved heterogeneity specific to each state and applies them to the analysis of labor market transitions. The proposed framework jointly estimates exits from employment and unemployment while approximating latent heterogeneity through a non-parametric bivariate discrete mixture in the spirit of Heckman and Singer (1984). In doing so, it allows for a richer representation of persistent unobserved differences across workers than standard single-state formulations.

The empirical application, based on a sample of young low-skilled workers in the Spanish labor market between 2000 and 2019, shows a clear asymmetry in the role of apprenticeship contracts. On the one hand, apprenticeship contracts are associated with lower separation rates during employment spells relative to other temporary contracts. On the other hand, once workers move into unemployment, a previous apprenticeship episode is not associated with faster transitions back into employment. Taken together, these results suggest that apprenticeship contracts may contribute to lower separation rates during employment spells, while their effects on subsequent re-employment prospects are more limited.

The results also show that unobserved heterogeneity is a central dimension of the transition process. The estimated bivariate discrete-mixture distribution reveals substantial differences across latent groups in their propensity to exit employment and unemployment, supporting the use of a framework in which heterogeneity is allowed to be state-specific and jointly distributed across both equations. In this sense, the model provides a more complete characterization of labor market dynamics than approaches based on separate single-state estimations.

From a practical point of view, the implementation of analytical expressions for the gradient vector and the Hessian matrix improves the numerical feasibility of estimating these models in large monthly-expanded longitudinal datasets. This computational gain is particularly valuable when richer latent structures are considered, since it makes an empirically relevant econometric specification operational for applied research.

Overall, the article makes both a methodological and an empirical contribution. Methodologically, it operationalizes the estimation of two-state hazard models with state-specific unobserved heterogeneity in a computationally efficient way. Empirically, it provides evidence on the distinct labor market trajectories associated with apprenticeship and temporary contracts in Spain. Future research may extend this framework to other labor market transitions, alternative institutional settings, and richer specifications of observed and unobserved heterogeneity. A further natural extension is to integrate this estimation framework more systematically into Stata–Python workflows for preprocessing, simulation, and post-estimation analysis in large-scale applications.

Author Contributions: Conceptualization, E.C. and D.T.-P.; methodology, N.R. and A.M.-K.; software, D.T.-P.; validation, E.C., D.T.-P., N.R. and A.M.-K.; formal analysis, E.C. and A.M.-K.; investigation, D.T.-P. and N.R.; resources, E.C.; data curation, A.M.-K.; writing—original draft preparation, E.C. and D.T.-P.; writing—review and editing, N.R. and A.M.-K.; visualization, D.T.-P.; supervision, E.C.; project administration, D.T.-P.; funding acquisition, E.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Interested readers can request the data and files needed to replicate the results presented in this work.

Acknowledgments: Authors acknowledge comments from participants at 2021 US Stata Conference and 2021 Canadian Stata Conference.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

Appendix A.1. Coefficients Estimates

Table A1. Exit from unemployment.

	Coeff.	Std. Error
Personal characteristics		
Female (=1)	−0.0438 ***	(0.00453)
Current age − 16	0.359 ***	(0.00623)
(Current age − 16) ²	−0.00769 ***	(0.000137)
Education: Compulsory stage #2	0.0130 **	(0.00509)
Not Spanish nationality (=1)	0.0432 ***	(0.00637)
Economic cycle		
Quarterly regional unemployment rate (Q.r.u.r.)	−0.0162 ***	(0.000311)
Regional fixed effects		
Andalucia	0.0414 ***	(0.00686)
Aragon	−0.0521 ***	(0.0124)
Asturias	−0.140 ***	(0.0182)
Baleares	−0.168 ***	(0.0147)
Canarias	−0.0640 ***	(0.0122)
Cantabria	−0.143 ***	(0.0233)
Castilla La Mancha	−0.0940 ***	(0.00937)
Castilla Leon	−0.0145	(0.00997)
Valencia	−0.0257 ***	(0.00753)
Extremadura	−0.0611 ***	(0.0138)
Galicia	−0.152 ***	(0.0112)
Murcia	0.0471 ***	(0.0105)
Navarra	−0.0454 ***	(0.0169)
Pais Vasco	−0.0206 *	(0.0117)
Rioja	−0.0198	(0.0214)
Unemployment duration effect		
Unemp.spell: Month 1 (=1)	1.181 ***	(0.00877)
Unemp.spell: Month 2 (=1)	1.381 ***	(0.00876)
Unemp.spell: Month 3 (=1)	1.032 ***	(0.00934)
Unemp.spell: Month 4 (=1)	0.828 ***	(0.00990)
Unemp.spell: Months 5 to 6 (=1)	0.586 ***	(0.00939)
Unemp.spell: Months 7 to 8 (=1)	0.508 ***	(0.00986)
Unemp.spell: Months 9 to 11 (=1)	0.868 ***	(0.00900)
Unemp.spell: Months 12 to 13 (=1)	0.813 ***	(0.0105)
Unemp.spell: Months 14 to 15 (=1)	0.260 ***	(0.0131)
Unemp.spell: Months 16 to 18 (=1)	0.201 ***	(0.0124)

Table A1. Cont.

	Coeff.	Std. Error
Unemp.spell: Months 19 to 22 (=1)	0.235 ***	(0.0120)
Unemp.spell: Months 23 to 24 (=1)	0.466 ***	(0.0152)
Previous job characteristics		
Previous job: Apprenticeship contract (=1)	−0.158 ***	(0.00637)
Economic sector of previous job:		
Manufacturing industry (=1)	−0.185 ***	(0.00685)
Economic sector of previous job:		
High qualified services (=1)	0.0385 ***	(0.00615)
Economic sector of previous job:		
Low qualified services (=1)	−0.00975 **	(0.00484)
Unobserved heterogeneity		
η_2^u	−0.978 ***	(0.00463)
η_2^e	−1.025 ***	(0.00343)
p_{11} (Prob. Type 1)	*** (9.89%)	(0.00165)
p_{12} (Prob. Type 2)	−0.118 *** (8.78%)	(0.00234)
p_{21} (Prob. Type 3)	0.548 *** (17.1%)	(0.00302)
p_{22} (Prob. Type 4)	1.871 *** (64.23%)	(0.00327)
Constant	−0.501 ***	(0.0325)
Observations		8,001,341
Unemployment observations		4,485,930
Log-likelihood		−2,549,148.4

$\eta_1^u = 0$; $\eta_1^e = 0$; Standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A2. Exit from employment.

	Coeff.	Std. Error
Personal characteristics		
Female (=1)	0.0418 ***	(0.00385)
Current age − 16	−0.0206 ***	(0.00472)
(Current age − 16) ²	−0.00104 ***	(0.000104)
Education: Compulsory stage #2	−0.136 ***	(0.00429)
Not Spanish nationality (=1)	0.159 ***	(0.00521)
Economic cycle		
Quarterly regional unemployment rate (Q.r.u.r.)	0.0160 ***	(0.000315)
(Q.r.u.r.) × Log(t)	0.00133 ***	(0.000480)
(Q.r.u.r.) × Log(t) ²	−0.00382 ***	(0.000181)
Regional fixed effects		
Andalucia	0.102 ***	(0.00580)
Aragon	0.127 ***	(0.0105)
Asturias	0.0638 ***	(0.0151)
Baleares	−0.00248	(0.0129)
Canarias	−0.0430 ***	(0.0106)
Cantabria	0.0561 ***	(0.0210)
Castilla La Mancha	0.235 ***	(0.00746)
Castilla Leon	0.116 ***	(0.00852)
Valencia	0.0629 ***	(0.00640)

Table A2. Cont.

	Coeff.	Std. Error
Extremadura	0.157 ***	(0.0113)
Galicia	0.0271 ***	(0.00907)
Murcia	0.106 ***	(0.00854)
Navarra	0.186 ***	(0.0142)
Pais Vasco	0.198 ***	(0.0100)
Rioja	0.276 ***	(0.0167)
Employment duration effect		
Emp.spell: Month 1 (=1)	1.640 ***	(0.0232)
Emp.spell: Months 2 to 3 (=1)	1.238 ***	(0.0246)
Emp.spell: Months 4 to 5 (=1)	1.037 ***	(0.0243)
Emp.spell: Month 6 (=1)	0.695 ***	(0.0246)
Emp.spell: Months 7 to 8 (=1)	1.001 ***	(0.0233)
Emp.spell: Month 9 (=1)	0.647 ***	(0.0243)
Emp.spell: Months 10 to 11 (=1)	0.641 ***	(0.0229)
Emp.spell: Month 12 (=1)	0.461 ***	(0.0251)
Emp.spell: Month 13 (=1)	1.147 ***	(0.0230)
Emp.spell: Months 14 to 17 (=1)	0.105 ***	(0.0221)
Emp.spell: Month 18 (=1)	0.0548 *	(0.0303)
Emp.spell: Months 19 to 22 (=1)	0.301 ***	(0.0217)
Emp.spell: Month 23 (=1)	−0.0410	(0.0358)
Emp.spell: Month 24 (=1)	0.112 ***	(0.0346)
Emp.spell: Months 25 to 29 (=1)	0.494 ***	(0.0211)
Emp.spell: Months 30 to 35 (=1)	−0.106 ***	(0.0249)
Emp.spell: Month 36 (=1)	−0.106 *	(0.0548)
Job characteristics		
Apprenticeship contract (=1)	−0.751 ***	(0.00775)
Economic sector of current job:		
Manufacturing industry (=1)	−0.494 ***	(0.00570)
Economic sector of current job:		
High qualified services (=1)	0.00393	(0.00503)
Economic sector of current job:		
Low qualified services (=1)	−0.136 ***	(0.00399)
Unobserved heterogeneity		
η_2^u	−0.978 ***	(0.00463)
η_2^e	−1.025 ***	(0.00343)
p_{11} (Prob. Type 1)	*** (9.89%)	(0.00165)
p_{12} (Prob. Type 2)	−0.118 *** (8.78%)	(0.00234)
p_{21} (Prob. Type 3)	0.548 *** (17.1%)	(0.00302)
p_{22} (Prob. Type 4)	1.871 *** (64.23%)	(0.00327)
Constant	−1.410 ***	(0.0332)
Observations		8,001,341
Employment observations		3,515,411
Log-likelihood		−2,549,148.4

$\eta_1^u = 0$; $\eta_1^e = 0$; Standard errors in parentheses; *** $p < 0.01$, * $p < 0.1$.

Appendix A.2. Descriptive Statistics

Table A3. Descriptive statistics. Employment state.

	Apprenticeship		Temporary	
	Mean	Std. Error	Mean	Std. Error
Personal characteristics				
Female	33.42%	(0.4717)	34.74%	(0.4761)
Current age – 16	2.81	(2.0909)	6.01	(3.2671)
(Current age – 16) ²	358.07	(82.7078)	494.91	(150.2143)
Education: Compulsory stage #2	39.09%	(0.4879)	26.11%	(0.4392)
Not Spanish nationality	5.66%	(0.2309)	16.26%	(0.3690)
Economic cycle				
Quarterly regional unemployment rate (Q.r.u.r.)	15.08%	(7.3161)	14.96%	(7.6004)
(Q.r.u.r.) × Log(t)	29.47	(22.3892)	19.44	(20.6556)
(Q.r.u.r.) × Log(t) ²	71.94	(71.5478)	43.72	(64.8598)
Regional covariates				
Andalucia	30.89%	(0.4620)	22.61%	(0.4183)
Aragon	1.53%	(0.1226)	3.01%	(0.1708)
Asturias	2.51%	(0.1564)	1.48%	(0.1210)
Baleares	1.40%	(0.1175)	2.20%	(0.1468)
Canarias	2.84%	(0.1660)	3.62%	(0.1868)
Cantabria	1.12%	(0.1052)	0.97%	(0.0980)
Castilla La Mancha	5.56%	(0.2292)	5.31%	(0.2244)
Castilla Leon	4.04%	(0.1968)	4.82%	(0.2142)
Valencia	9.90%	(0.2987)	10.94%	(0.3121)
Extremadura	5.06%	(0.2192)	2.19%	(0.1464)
Galicia	10.00%	(0.3000)	4.72%	(0.2122)
Murcia	3.67%	(0.1880)	4.57%	(0.2088)
Navarra	0.79%	(0.0887)	1.63%	(0.1268)
Pais Vasco	2.05%	(0.1418)	3.77%	(0.1906)
Rioja	0.18%	(0.0423)	0.83%	(0.0907)
Employment duration				
Emp.spell: Month 1	9.14%	(0.2881)	26.65%	(0.4421)
Emp.spell: Months 2 to 3	15.59%	(0.3627)	24.77%	(0.4317)
Emp.spell: Months 4 to 5	12.64%	(0.3322)	12.91%	(0.3353)
Emp.spell: Month 6	5.73%	(0.2323)	4.51%	(0.2075)
Emp.spell: Months 7 to 8	9.66%	(0.2954)	6.73%	(0.2505)
Emp.spell: Month 9	4.00%	(0.1960)	2.51%	(0.1565)
Emp.spell: Months 10 to 11	7.45%	(0.2626)	3.98%	(0.1955)
Emp.spell: Month 12	3.50%	(0.1838)	1.60%	(0.1255)
Emp.spell: Month 13	3.38%	(0.1806)	1.41%	(0.1180)
Emp.spell: Months 14 to 17	10.50%	(0.3065)	3.96%	(0.1951)
Emp.spell: Month 18	2.38%	(0.1523)	0.79%	(0.0890)
Emp.spell: Months 19 to 22	7.03%	(0.2556)	2.56%	(0.1581)
Emp.spell: Month 23	1.48%	(0.1207)	0.52%	(0.0720)
Emp.spell: Month 24	1.43%	(0.1189)	0.48%	(0.0696)
Emp.spell: Months 25 to 29	3.03%	(0.1713)	1.96%	(0.1389)
Emp.spell: Months 30 to 35	2.05%	(0.1416)	1.69%	(0.1290)
Emp.spell: Month 36	0.30%	(0.0548)	0.23%	(0.0480)
Job characteristics				
Economic sector of current job:				
Manufacturing industry	16.30%	(0.3693)	11.56%	(0.3198)

Table A3. Cont.

	Apprenticeship		Temporary	
	Mean	Std. Error	Mean	Std. Error
Economic sector of current job:				
High qualified services	3.56%	(0.1853)	09.75%	(0.2966)
Economic sector of current job:				
Low qualified services	13.16%	(0.3381)	18.80%	(0.3907)
Observations	260,826		2,164,478	

Table A4. Descriptive statistics. Unemployment state.

	Apprenticeship		Temporary	
	Mean	Std. Error	Mean	Std. Error
Personal characteristics				
Female	37.16%	(0.4832)	37.86%	(0.4850)
Current age – 16	5.17	(2.7069)	5.93	(3.2998)
(Current age – 16) ²	455.58	(119.8669)	491.88	(152.7011)
Education: Compulsory stage #2	33.81%	(0.4730)	22.47%	(0.4174)
Not Spanish nationality	4.75%	(0.2128)	15.83%	(0.3651)
Economic cycle				
Quarterly regional unemployment rate (Q.r.u.r.)	16.67%	(8.0199)	16.53%	(8.0060)
Regional covariates				
Andalucia	28.28%	(0.4503)	22.04%	(0.4145)
Aragon	1.64%	(0.1270)	3.12%	(0.1740)
Asturias	2.66%	(0.1612)	1.22%	(0.1098)
Baleares	1.87%	(0.1356)	2.53%	(0.1572)
Canarias	3.18%	(0.1754)	3.82%	(0.1918)
Cantabria	1.09%	(0.1039)	0.79%	(0.0887)
Castilla La Mancha	5.27%	(0.2235)	5.95%	(0.2366)
Castilla Leon	4.50%	(0.2075)	5.09%	(0.2199)
Valencia	10.10%	(0.3014)	11.19%	(0.3153)
Extremadura	4.22%	(0.2011)	2.61%	(0.1595)
Galicia	9.73%	(0.2964)	3.67%	(0.1881)
Murcia	3.48%	(0.1833)	4.53%	(0.2081)
Navarra	0.95%	(0.0973)	1.59%	(0.1251)
Pais Vasco	2.27%	(0.1491)	3.46%	(0.1828)
Rioja	0.36%	(0.0605)	0.94%	(0.0965)
Unemployment duration				
Unemp.spell: Month 1	9.58%	(0.2943)	11.01%	(0.3129)
Unemp.spell: Month 2	8.19%	(0.2743)	9.25%	(0.2897)
Unemp.spell: Month 3	6.77%	(0.2513)	7.55%	(0.2642)
Unemp.spell: Month 4	5.92%	(0.2360)	6.55%	(0.2474)
Unemp.spell: Months 5 to 6	10.10%	(0.3014)	11.09%	(0.3141)
Unemp.spell: Months 7 to 8	8.38%	(0.2771)	9.08%	(0.2873)
Unemp.spell: Months 9 to 11	10.07%	(0.3010)	10.61%	(0.3080)
Unemp.spell: Months 12 to 13	5.21%	(0.2224)	4.94%	(0.2168)
Unemp.spell: Months 14 to 15	4.48%	(0.2070)	4.01%	(0.1961)
Unemp.spell: Months 16 to 18	5.74%	(0.2327)	5.07%	(0.2194)
Unemp.spell: Months 19 to 22	6.24%	(0.24189)	5.44%	(0.2269)
Unemp.spell: Months 23 to 24	2.61%	(0.15950)	2.24%	(0.1482)
Previous job characteristics				
Economic sector of previous job:				
Manufacturing industry	12.92%	(0.3354)	9.46%	(0.2926)

Table A4. Cont.

	Apprenticeship		Temporary	
	Mean	Std. Error	Mean	Std. Error
Economic sector of previous job:				
High qualified services	7.04%	(0.2559)	9.93%	(0.2991)
Economic sector of previous job:				
Low qualified services	19.26%	(0.3943)	20.24%	(0.4018)
Observations	698,772		2,164,478	

References

- Ahn, H. J. (2023). The role of observed and unobserved heterogeneity in the duration of unemployment. *Journal of Applied Econometrics*, 38(1), 3–23. [\[CrossRef\]](#)
- Alba-Ramirez, A., Arranz, J. M., & Muñoz-Bullón, F. (2007). Exits from unemployment: Recall or new job. *Labour Economics*, 14, 788–810. [\[CrossRef\]](#)
- Alba-Ramirez, A., Arranz, J. M., & Muñoz-Bullón, F. (2012). Re-employment probabilities of unemployment benefit recipients. *Applied Economics*, 44(28), 3645–3664. [\[CrossRef\]](#)
- Allison, P. D. (1982). Discrete-time methods for the analysis of event histories. *Sociological Methodology*, 13, 61–98. [\[CrossRef\]](#)
- Arranz, J. M., & García-Serrano, C. (2011). Are the MCVL tax data useful? Ideas for mining. *Hacienda Pública Española*, 199(4), 151–186.
- Arranz, J. M., & García-Serrano, C. (2014). The interplay of the unemployment compensation system, fixed-term contracts and rehiring: The case of Spain. *International Journal of Manpower*, 35(8), 1236–1259. [\[CrossRef\]](#)
- Arranz, J. M., García-Serrano, C., & Hernanz, V. (2013). How do we pursue labormetrics? An application using the MCVL. *Estadística Española*, 55(181), 231–254.
- Arranz, J. M., García-Serrano, C., & Toharia, L. (2010). The influence of temporary employment on unemployment exits in a competing risk framework. *Journal of Labor Research*, 31, 67–90. [\[CrossRef\]](#)
- Bentolila, S., García Pérez, J. I., & Jansen, M. (2017). Are the Spanish long-term unemployed unemployable? *SERIEs—Journal of the Spanish Economic Association*, 8, 1–41.
- Bieszk-Stolorz, B. (2021). Models of competing events in assessing the effects of the transition of unemployed people between the states of registration and de-registration. In K. Jajuga, K. Najman, & M. Walesiak (Eds.), *Data analysis and classification. SKAD 2020* (pp. 213–228). Springer. [\[CrossRef\]](#)
- Carrasco, R., Gálvez-Iniesta, I., & Jerez, B. (2024). Do temporary help agencies help? Employment transitions for low-skilled workers. *Labour Economics*, 90, 102586. [\[CrossRef\]](#)
- Castro, R., Lange, F., & Poschke, M. (2024). Labor force transitions. In *NBER working paper 33200*. National Bureau of Economic Research. [\[CrossRef\]](#)
- Cueto, B., & López, F. (2019). The apprenticeship contract: An evaluation. *Review of Public Economics*, 231, 15–39. [\[CrossRef\]](#)
- Eberlein, L., Garnier-Villarreal, M., & Pavlopoulos, D. (2024). Starting flexible, always flexible? The relation of early temporary employment and young workers' employment trajectories in the Netherlands. *Research in Social Stratification and Mobility*, 89, 100861. [\[CrossRef\]](#)
- Felgueroso, F., García Pérez J. I., Jansen, M., & Troncoso-Ponce, D. (2018). The surge in short-duration contracts in Spain. *De Economist*, 166, 503–534. [\[CrossRef\]](#)
- García-Pérez, J. I. (2008). La Muestra Continua de Vidas Laborales: Una guía de uso para el análisis de transiciones. *Revista de Economía Aplicada*, 16(E-1), 5–28.
- Gaure, S., Røed, K., & Zhang, T. (2007). Time and causality: A Monte Carlo assessment of the timing-of-events approach. *Journal of Econometrics*, 141, 1159–1195. [\[CrossRef\]](#)
- Gould, W., Pitblado, J., & Poi, B. (2010). *Maximum likelihood estimation with Stata*. Stata Press.
- Heckman, J. J., & Singer, B. (1984). A method for minimizing the impact of the distributional assumptions in econometric models for duration data. *Econometrica*, 52, 271–320. [\[CrossRef\]](#)
- Jansen, M., & Troncoso-Ponce, D. (2018). The impact of apprenticeships on the labour market insertion of youth in Spain. In *New skills at work*. FEDEA & J.P. Morgan.
- Jenkins, S. (1995). Easy estimation methods for discrete-time duration models. *Oxford Bulletin of Economics and Statistics*, 57(1), 129–136. [\[CrossRef\]](#)
- Kyyra, T., Arranz, J. M., & García-Serrano, C. (2019). Does subsidized part-time employment help unemployed workers to find full-time employment? *Labour Economics*, 56, 68–83. [\[CrossRef\]](#)
- Lafuente, C. (2020). Unemployment in administrative data using survey data as a benchmark. *SERIEs—Journal of the Spanish Economic Association*, 11(2), 115–153. [\[CrossRef\]](#)

- Lancaster, T. (1992). *The econometric analysis of transition data*. Cambridge University Press.
- Lapuerta, I. (2010). *Claves para el trabajo con la muestra continua de vidas laborales*. DemoSoc working paper. Universitat Pompeu Fabra.
- Lechner, M., & Pfeiffer, F. (2001). *Econometric evaluation of labour market policies*. Springer Science & Business Media.
- Machado, R. J. M., & van den Hout, A. (2018). Flexible multistate models for interval-censored data: Specification, estimation, and an application to ageing research. *Statistics in Medicine*, 37(10), 1636–1649. [CrossRef]
- Nicoletti, C., & Rondinelli, C. (2010). The (mis)specification of discrete duration models with unobserved heterogeneity: A Monte Carlo study. *Journal of Econometrics*, 159, 1–13. [CrossRef]
- Rebollo-Sanz, Y. F., & García-Pérez, J.I. (2015). Are unemployment benefits harmful to the stability of working careers? The case of Spain. *SERIEs—Journal of the Spanish Economic Association*, 6, 1–41. [CrossRef]
- Troncoso-Ponce, D. (2017). *Faster estimation of discrete time duration models with unobserved heterogeneity using hshaz2*. Working paper. Universidad Pablo de Olavide, Department of Economics. Available online: <https://ideas.repec.org/p/pab/wpaper/17.05.html> (accessed on 15 April 2026).
- Troncoso-Ponce, D. (2018). *Estimation of competing risks duration models with unobserved heterogeneity using hsmlogit*. Working paper. Universidad Pablo de Olavide. Available online: <https://www.upo.es/serv/bib/wps/econ1803.pdf> (accessed on 15 April 2026).
- van den Hout, A., & Tan, W. (2019). Flexible parametric multistate modelling of employment history. *Statistical Modelling*, 19(3), 323–338. [CrossRef]
- Vaquero García, A., Cruz González, M. M., & Suárez Porto, V. M. (2024). Evaluation of training contracts for young people in Spain, Germany, and France. *Employee Responsibilities and Rights Journal*, 36(3), 385–399. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.