

Application of latent growth modeling to identify different working life trajectories: the case of the Spanish WORKss cohort

by Laura Serra, PhD,^{1,2} María Andrée López Gómez, MPH,^{1,2,3} Albert Sanchez-Niubo, PhD,^{1,2,3} George L Delclos, MD,^{1,2,3,4} Fernando G Benavides, PhD^{1,2,3}

Serra L, López Gómez MA, Sanchez-Niubo A, Delclos GL, Benavides FG. Application of latent growth modeling to identify different working life trajectories: the case of Spanish WORKss cohort. *Scand J Work Environ Health* – online first. doi:10.5271/sjweh.3606

Objective The aim of this study was to describe the application of latent class growth analysis (LCGA) to identify different working life trajectories (WLT) using employed working time by year as a repeated measure.

Methods Trajectories are estimated using LCGA, which considers all individuals within a trajectory to be homogeneous. The methodology was applied to a subsample of the Spanish WORKing life Social Security (WORKss) cohort, limited to persons born 1956–1965 (N=247 475). The number of days worked per year is used as a repeated measure across 32 time points (1981–2013).

Results According to the model-fit results and further guided by expert knowledge, a four WTL model was selected as the optimal approach: WLT1 or “high labor force participation” (N=99 591; 40.2%); WLT2 or “decreased labor force participation” (N= 22 846; 9.2%); WLT3 or “increased labor force participation” (N=59 213; 23.9%); and WLT4 or “low labor force participation” (N=65 827; 26.6%). WLT1 consisted mainly of men with more years of work experience (>19 years) while WLT4 was mainly composed by women with <9 years. The other two trajectories had opposite trends and no sex differences. The occupational category variable had little influence in the trajectories.

Conclusions Longitudinal data that are regularly collected by administrative systems can benefit from LCGA approaches to identify different trajectory patterns that may be associated with an outcome of interest. In occupational epidemiology, this study represents a step forward by using this modeling approach to identify different WLT.

Key terms cohort study; employment condition; labor market; latent class growth analysis; latent growth modeling; occupational health; Spain.

In occupational epidemiology, there is significant interest in analyzing causal relationships between work-related factors, socioeconomic status, and health. This interplay has a dynamic relationship during the life course of a worker (1, 2). In that sense, in a changing labor market, cross-sectional designs fall short because they capture health and work status only at one moment in time and do not provide any information on previous work life and health status. Researchers with access to repeated measures should explore which methodologies

can adequately address research questions that include time-varying data.

Recently, computational advances and the large amount of data that administrative systems are continuously collecting allow the use of statistical techniques that have the potential to obtain more interesting and reliable results. The Spanish social security system collects longitudinal information related to work and pensions, dating back to 1981. This information includes repeated measures of variables with exact dates in time,

¹ Center for Research in Occupational Health (CiSAL), Department of Experimental and Health Sciences, Pompeu Fabra University, Barcelona, Spain.

² IMIM Parc Salut Mar, Social epidemiology and occupational health group, Barcelona, Spain.

³ CIBER of Epidemiology and Public Health, Spain.

⁴ Department of Epidemiology, Human Genetics and Environmental Sciences, The University of Texas School of Public Health (UT), Houston, USA.

Correspondence to: Albert Sanchez-Niubo, Center for Research in Occupational Health (CiSAL), Department of Experimental and Health Sciences, Pompeu Fabra University, Barcelona, Spain. [E-mail: albert.sanchez@pssjd.org]

which is useful for research on the role of employment and employment-related events over the life course and their relation to worker health. More specifically, access to repeated measures allows the construction of working life trajectories (WLT) by one key variable, such as employment time, allowing us to examine the pathway an individual follows over a lifetime. This approach is relatively novel in occupational health research. At a population level, we are then able to group individuals with similar WLT, enabling the study of the associations between different trajectories, other determinants like socioeconomic status (SES) and outcomes such as permanent disability or mortality. Although there are several statistical approaches to obtain heterogeneous trajectories, our work focuses on one of them, namely latent growth modeling (LGM) (3, 4).

LGM techniques represent a flexible statistical approach for longitudinal data to capture information about interindividual differences in intraindividual changes over time taking into account unobserved heterogeneity within a larger population (3, 5, 6). This allows us to represent unobserved heterogeneity by means of separate growth models for each latent class (trajectory), each with its unique estimates of variances and covariate influences. There are two similar LGM approaches to this end: growth mixture modeling (GMM) (7, 8) and latent class growth analysis (LCGA) (9). GMM can be described as a multilevel, random-effects model that allows for variability in the growth factors within each class (variance and covariance). On the other hand, LCGA is a special type of GMM in which random effects (variance and covariance) for the growth parameters within each trajectory are assumed to be fixed at zero, and considers all individuals in a trajectory as homogeneous (10, 11, 12). In this sense, GMM depicts more accuracy in terms of discriminating individuals within a trajectory. However, large datasets may benefit more from LCGA as convergence is more stable and requires a lower computational workload. This feature makes LCGA more appropriate when dealing with complex models (3, 4).

The application of LGM in the field of occupational health is currently growing and becoming popular (1). These approaches have mostly been used to capture unknown heterogeneity in work-related factors and to see their association with some health outcomes. For example, to study the association between occupational complexity and cognition in later life (13), work characteristics and hypertension (14), absenteeism trajectories and health factors (15), or work functioning and return to work after sick leave (16), among other similar studies (17, 18, 19, 20). In summary, these approaches allow us to address labor market changes from a life course perspective (21).

The aim of this study was to describe the applica-

tion of LCGA to identify different WLT using employed working time by year as a repeated measure, taking advantage of the Spanish WORKing life Social Security (WORKss) cohort (22).

Methods

Data

The WORKss cohort is built on an annual random sample of 4% of individuals registered in the Spanish social security system (through contributions or pension) from 2004 to 2013 (22). The dataset provides information about the exact dates (day/month/year) of any contact with social security, allowing the reconstruction of complete labor trajectories from 1981 (the first year that the Spanish social security data is considered to be exhaustive and of high quality) to 2013 (the last available year).

We focus on a subsample of the WORKss cohort corresponding to individuals born 1956–1965 because they have had the opportunity of a longer work trajectory. In 1981, these individuals were 16 (the start of working age in Spain) to 25 years old, and in 2013 they were 48–57 years old. The number of individuals considered was 247 477 (45.5% females).

We used the number of days worked per year (from 1981–2013) as independent variable. Main individual-level variables included in this dataset were: sex, age, years worked, and occupational category. The variable years worked is the sum of the time a person has been working in any kind of contract from 1981–2013 and converting it to years. Occupational category is built from the contribution rate groups reported by the employer, which considers the physical demands and qualifications required for a job (23). This variable includes four categories: skilled non-manual, skilled manual, unskilled non-manual, and unskilled manual. Since individuals may have several occupations over a lifetime, the occupational category assigned in our study was defined by the longest held occupation over this period.

In order to better understand the different profiles assigned to each individual, the identified trajectories were further described according to variables of interest (sex, mean age, number of years worked, and occupational category).

Statistical analysis

To construct the WLT, we built 32 measurement periods for each participant, one per year of observation (1981–2013) containing the days worked. We introduced zero working days if the person was inactive during a certain year but had an administrative activity before

and after that period. If individuals were not working a certain year due to permanent disability, early retirement or death, a missing value is coded because in these situations the individual is not able to work. We used LCGA to identify different WLT.

Trajectories can be estimated assuming they follow any equation, either a linear function, a quadratic function or any expression including more complex elements. The number of these parameters, called latent growth factors, can vary according to the data, but it is recommended to apply the criterion of parsimony and use the fewest number of parameters to obtain similar results. Figure 1 features the LGM representation consisting of j latent growth trajectories with three latent growth factors, intercept (I_j), slope (S_j) and the quadratic growth factor (Q_j), formed by the observed variables that represent repeated measures across 32 time points (1981–2013).

As individuals are assigned in the different trajectories according to the highest probability of belonging to each one, trajectories can also be constructed calculating the mean of yearly worked days for all the individuals within each trajectory. So two types of trajectories can be constructed: on the one hand, the estimated WLT by means of the latent growth factors, intercept (I) and slope (S) in this case and, on the other hand, the observed WLT by calculating the mean of days worked. Both are presented so that the fit of the estimation can be compared to the observed values.

There are different criteria for assessing the optimal number of latent trajectories. On the one hand, an initial approach can be based on prior knowledge of the researcher. On the other hand, a combination of fit indices can be considered, such as the entropy (near to 1), a significant Lo, Mendell and Rubin likelihood ratio test (LMR-LRT) or a significant bootstrap likelihood ratio test (BLRT). The LMR-LRT and BLRT tests provide information to avoid overestimation of the number of trajectories. In particular, these two tests allow for the comparison of model fit between subsequent models,

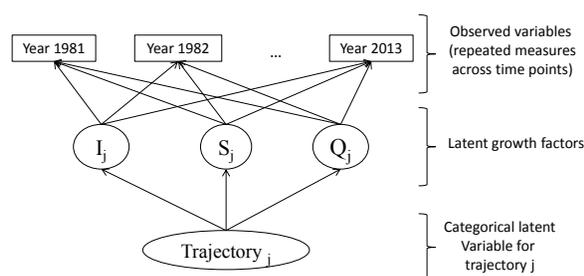


Figure 1. Representation of a latent class growth analysis assuming three latent growth factors: intercept I_j , slope S_j and quadratic element Q_j .

and then a significant P-value suggests that a model with one more trajectory provides better results in terms of goodness of fit. Model fit can be also assessed with the Akaike information criterion (AIC), Bayesian information criterion (BIC) and the sample-size-adjusted BIC, with the lowest values indicating the best model (24). In addition, other elements can also help specify the number of trajectories, such as parsimony, a theoretical justification, successful convergence, high posterior probabilities (near 1.0) or interpretability (eg, classifying <1% of individuals in one class may make no sense) (25, 26, 27).

Results

We began the analysis using two latent growth factors (the intercept and the slope) and then repeated our analysis also considering the quadratic factor. The representation of the trajectories was nearly identical, so we selected the more parsimonious approach using only the two latent growth factors (the corresponding A in table 1).

Then, after testing different models with these two latent growth factors (group A of WLT), we began building a model with two trajectories, then three trajectories and so on until an optimal number of trajectories was identified. The model with at ≥ 5 WLT was identified as the probably best solution according to the fit indices. The AIC and BIC indices were the smallest, the LMR-LRT and BLRT indices were statistically significant regarding to subsequent models and, lastly, the entropy index was >0.95 , as in the other models. However, when comparing the graphical representation using WLT4 and WLT5 (figure 2), the additional WLT (WLT5 in the bottom figure 2) shows a slope similar to that of WLT3. Therefore, we decided to select WLT4 because we felt WLT5 did not provide enough information to sufficiently justify its inclusion, based on our previous knowledge of the Spanish labor market.

The four trajectory models were roughly classified as: WLT1, people who were employed most of the time ($N=99\ 591$; 40.2%), or "high labor force participation"; WLT2, people with an average number of worked days (around 200 days per year) including a peak around 1989 and decreasing afterwards, or "decreased labor force participation" ($N=22\ 846$; 9.2%); WLT3, people who worked for a few days in the 1980s but increased the days worked during the rest of the study period ($N=59\ 213$; 23.9%), classified as "increased labor force participation"; and WLT4, people who mostly did not work during the analyzed years except for an increase around 2005 ($N=65\ 827$; 26.6%), or "low labor force participation".

Table 1. Model-fit results for latent class growth analysis specifying two groups of working life trajectories (WLT) according to (A) two latent growth factors (intercept and slope) and (B) three latent growth factors (intercept, slope and quadratic). [AIC=Akaike information criterion; BIC=Bayesian information criterion; LMR-LRT=Lo, Mendell & Rubin likelihood ratio test; BLRT=bootstrap likelihood ratio test.]

WLT number		AIC ^a	BIC ^b	Sample-size-adjusted BIC ^b	LMR-LRT ^c P-value	BLRT ^d P-value	Entropy
A	2	96 288 015	96 288 390	96 288 275	<0.001	<0.001	0.982
	3	95 487 681	95 488 088	95 487 963	<0.001	<0.001	0.961
	4	94 980 231	94 980 668	94 980 534	<0.001	<0.001	0.968
	5	94 608 732	94 609 201	94 609 057	<0.001	<0.001	0.957
B	3	95 004 630	95 005 068	95 004 934	<0.001	<0.001	0.968
	4	94 452 609	94 453 088	94 452 941	<0.001	<0.001	0.960
	5	93 943 179	93 943 700	93 943 541	<0.001	<0.001	0.966

Table 2 describes the four trajectories considering sex, mean age, number of years worked, occupational category. Regarding sex, the decreased and increased labor force participation trajectories contained roughly the same proportion of men and women and had more or less the same distribution of people in each category of years worked. Men (73%) dominated in the high labor force participation and women (63%) dominated in the low labor force participation trajectory. As expected, the high labor force participation trajectory mostly included individuals with >19 years of work experience, whereas the opposite occurred in the low labor force participation trajectory. When occupational category was considered, no such differences between trajectories were present, except for increased and low labor force participation, both of which included a lower percentage of non-manual skilled workers.

Discussion

LGM approaches are a powerful statistical tool for identifying different latent trajectories, which in turn allow the classification of individuals into specific trajectories. From a life course perspective, this categorization enables the identification of different patterns shaped by working lives over time rather than focusing in cross-sectional studies. Thus, taking advantage of longitudinal data on working people, LGM approaches allow the identification of different labor paths rather than looking at individual evolution over time. In this study, we focused on LCGA and applied it to a subsample of the Spanish WORKss cohort, which provides data based on a 32-year follow-up period. Before LCGA we used GMM, but we encountered convergence problems due to the high number of parameters based on 32 time points (28).

Besides identifying different trajectories, LCGA also assigns each individual to the WLT to which they have the highest probability of belonging. This classification produces a new variable that was not so easily managed before applying the LCGA. This new variable can then

be included in a subsequent analysis as a categorical explanatory variable where each trajectory would be a category of the variable. The idea is to assign a latent class to each individual in a sample, adding a new variable to the original dataset that incorporates latent information not previously available.

WLT were summarized into four categories based on differences in their slopes. The “high labor force participation” trajectory was the one with the greatest number of worked days and a low positive slope and included the largest percentage of individuals in the sample (40.2%), most of whom were men. This trajectory could be described as the one with a stable working life trajectory because individuals had been working most of the time. In fact, the trajectory with the lowest percentage (9.2%) corresponded to the “decreased labor force participation”, which was the only one with a negative slope and composed mainly of women. Differences regarding sex could be explained by the fact that our study sample represents a period of time where women had a low participation rate in the Spanish labor market. Their participation increased at the end of our follow-up period (29). The two remaining trajectories, with small differences regarding sex, showed positive slopes that had similar starting values; however, one increased much faster than the other one. Finally, it is worth mentioning that the occupational category did not have a strong influence in the results. This may be due to the period covered by the data; a period of economic expansion which brought a wide demand of all different occupational categories. These results showed a positive trend as most of individuals in the sample, about 90% of workers included in these trajectories increased the days worked per year as time elapsed.

The decision about the number of trajectories was based, in the first and main step, on the classification accuracy values, but also, in a second and complementary step, on the representativeness and distinctiveness of the trajectories, driven by the knowledge and experience of the researcher. Combining statistical criteria with expert knowledge was felt to be the best way to identify the optimal model.

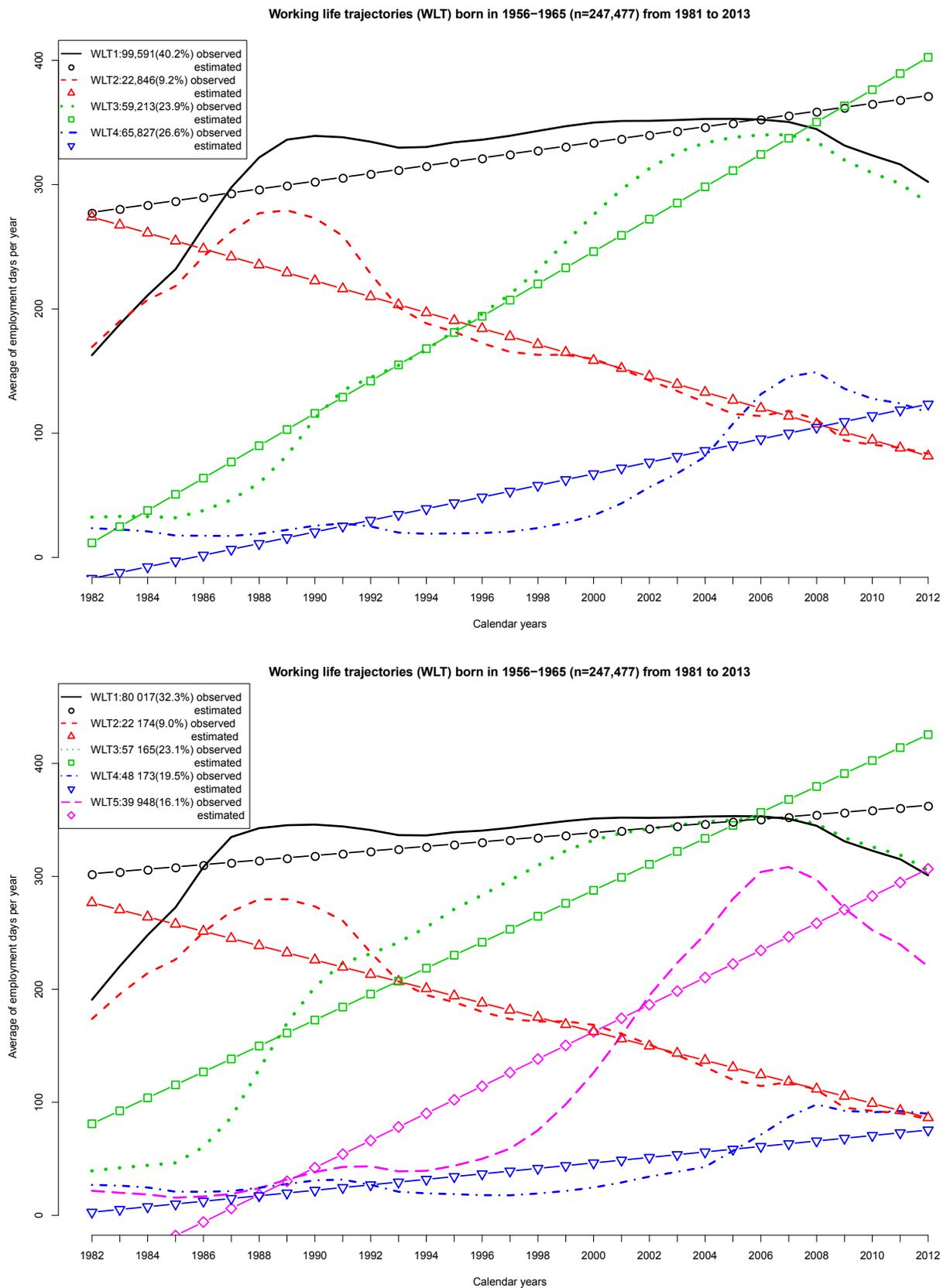


Figure 2. Observed and estimated growth trajectories considering the annual number of worked days (1981–2013) among people from the subsample of the Spanish WORKing life Social Security (WORKss) cohort, limited to individuals born 1956–1965 (N= 247 477). On the top, four working life trajectories (WLT) are represented and on the bottom, five WLT.

The challenge here was to classify individuals with similar workloads over time into different groups, resulting in a new categorical variable that we called WLT. In our study, these LCGA-identified WLT were explicitly based on the days worked each year. However, other WLT could have been identified if we had used other variables included in the dataset – such as the number of contracts or the time in unemployment – as the independent variable. Furthermore, other variables not included in the dataset (eg, the number of children) or other information that we are not even aware of (eg, the distance between place of residence and work) could influence the heterogeneity identified by LCGA. Certainly, these unknown variables can be relevant if they are associated with health outcomes of interest according to our hypothesis.

The LGM literature includes several studies that have used the methodology described in this paper for the analysis of repeated measures data for the purpose of identifying individual trajectories given certain variables (8, 13, 14, 30, 31). In an occupational study, Raymo et al (2) identified latent class employment trajectories using the probability of being employed; the trajectory patterns they identified were quite similar to ours, but none centered on the analysis of WLT as we have done. To our knowledge, this is the first study focused on the construction of WLT. In particular, we identified different working latent trajectories over time considering the average number of worked days per year in a representative sample of a country workforce. In addition, this study was based on a large dynamic cohort with a large number of independent measures.

Our dataset presents a left-sided truncation as we

only considered persons who had contact with the social security system in 2004–2013. We were not able to collect data on those who did not have contact within this period. Thus, for example, the last part of the WLT4 might have been flatter if we had been able to observe people in contact with the social security system prior to 2004. It is likely that these persons did not work during the period 2004–2013 and, thus, the average number of days worked per year would have decreased and the trajectory would have been lower.

Another limitation is the absence of information on informal employment, defined as workers lacking social security protections (32). Thus, the WLT we identified only represent those Spanish workers who had a formal contract and were eligible for social protection. Moreover, we estimated the trajectories based only on the number of days worked per year. Other factors, such as the number of new entries into the workforce or annual number of contracts, could also be interesting for classifying each individual as a reflection of the flexibility experienced in trajectories. For example, it is not the same thing to be employed for a similar number of days with one versus four contracts.

Longitudinal data that are regularly collected by administrative systems can benefit from LGM approaches to identify different trajectory patterns that may be associated with an outcome of interest. In occupational epidemiology, this study provides a step forward by using this modeling approach to identify different WLT. Examining the effect of these WLT on health outcomes may be useful in the design of prevention policies. However, there is a dynamic relationship between the heterogeneity identified by LGM and many

Table 2. Description of the four working life trajectories (WLT) based on two latent growth factors (intercept and slope) and according to some covariates. [SD=standard deviation].

Covariates	WLT1: high labor force participation (N=99 591)				WLT2: decreased labor force participation (N=22 846)				WLT3: increased labor force participation (N=59 213)				WLT4: low labor force participation (N=65 827)			
	N	%	Mean	SD	N	%	Mean	SD	N	%	Mean	SD	N	%	Mean	SD
Sex																
Men	72 550	72.8			11 699	51.2			26 312	44.4			24 353	37.0		
Women	27 041	27.2			11 147	48.8			32 901	55.6			41 474	63.0		
Age			36.76	2.57			36.78	2.58			36.78	2.57			36.74	2.56
Years worked (in any kind of contract)																
0–4	0	0			1	0			0	0			35 901	54.5		
5–9	0	0			3358	14.7			1974	3.3			27 327	41.5		
10–14	0	0			8690	38.0			17 276	29.2			2599	3.9		
15–19	412	0.4			6939	30.4			17 121	28.9			0	0		
>19	99 179	99.6			3858	16.9			22 842	38.6			0	0		
Occupational category																
Non-manual and skilled	16 172	16.3			1033	5.0			9276	15.7			2881	4.5		
Non-manual and unskilled	25 592	25.8			4643	22.7			14 223	24.1			9418	14.7		
Manual and skilled	27 526	27.7			6686	32.7			11 669	19.8			14 772	23.0		
Manual and unskilled	9883	10.0			5089	24.9			9945	16.9			21 098	32.9		
Non-informed	20 076	20.2			3013	14.7			13 832	23.5			15 945	24.9		

known and unknown time-varying determinants that is not easy to interpret. Accessing information of work, socioeconomic status and health on the same individual over time is of paramount importance to control, as much as possible, whether statistical effects are due to causality or bias (33). More complex LGM is needed to more properly analyze the issues this study raises.

Declarations

Ethics approval and consent to participate were not required because the Spanish Ministry of Employment and Social Security provided the data and had already controlled the ethical process.

Conflict of interests

The authors declare no conflicts of interest.

Funding

This work was supported by grants from the Instituto Salud Carlos III-FEDER (FIS PI14/00057), the Spanish National Health Institute Carlos III (FIS 08/0914 and FIS 11/01470), and the CIBER of Epidemiology and Public Health. The funders had no involvement in the study design, the collection, analysis and interpretation of data, the writing of the report, or the decision to submit the paper for publication.

Acknowledgements

The authors would like to thank the General Directorate for the Organization of the Social Security (Spanish Ministry of Employment and Social Security) for providing data of the Continuous Working Life Sample (CWLS), the data source of the WORKss cohort. The authors would also like to thank Xavier Durán for his support on managing the dataset and contributing to the construction of the WORKss cohort.

References

- Burdorf A. The need for novel strategies to analyze the dynamic pattern of worker's health over time and the consequences for sustained employability. *Scand J Work Environ Health*. 2012;38:485–8. <https://doi.org/10.5271/jweh.3325>.
- JM Raymo AHM, Warren J. Careers and Mortality: What Do We Learn from Detailed Employment History Data? University of Wisconsin-Madison; 2013. Working paper.
- Jung T, Wickrama KAS. An Introduction to Latent Class Growth Analysis and Growth Mixture Modeling *Social Pers Psych Compass*. 2008;2(1):302–17.
- Li F, Duncan TE, Duncan SC, Acock A. Latent Growth Modeling of Longitudinal Data: A Finite Growth Mixture Modeling Approach. *Structural Equation Modeling: A Multidisciplinary Journal*. 2001;8(4):493–530. https://doi.org/10.1207/S15328007SEM0804_01.
- Berlin KS, Parra GR, Williams NA. An Introduction to Latent Variable Mixture Modeling (Part 2): Longitudinal Latent Class Growth Analysis and Growth Mixture Models. *Journal of Pediatric Psychology*. 2013;39(2):188–203. <https://doi.org/10.1093/jpepsy/jst085>.
- Green M. Latent class analysis was accurate but sensitive in data simulations. *J Clin Epidemiol*. 2014;67:1157–62. <https://doi.org/10.1016/j.jclinepi.2014.05.005>.
- Wang M, Bodner TE. Growth Mixture Modeling: Identifying and Predicting Unobserved Subpopulations With Longitudinal Data. *Organizational Research Methods*. 2007;10(4):635–56. <https://doi.org/10.1177/1094428106289397>.
- Muthén B, Muthén L. Integrating person-centered and variable-centered analyses: growth mixture modeling with latent trajectory classes. *Alcohol Clin Exp Res*. 2000 Jun;24:882–91. <https://doi.org/10.1111/j.1530-0277.2000.tb02070.x>.
- Nagin DS, Land KC. Age criminal careers and population heterogeneity: specification and estimation of a nonparametric mixed Poisson model. *Criminology*. 1993 aug;31(3):327–62. <https://doi.org/10.1111/j.1745-9125.1993.tb01133.x>.
- Nagin DS. Analyzing developmental trajectories: A semiparametric group-based approach. *Psychological Methods*. 1999;4(2):139–57. <https://doi.org/10.1037/1082-989X.4.2.139>.
- Roeder K, Lynch KG, Nagin DS. Modeling Uncertainty in Latent Class Membership: A Case Study in Criminology. *Journal of the American Statistical Association*. 1999;94(447):766–76. <https://doi.org/10.1080/01621459.1999.10474179>.
- Kreuter F, Muthén B. Longitudinal modeling of population heterogeneity: methodological challenges to the analysis of empirically derived criminal trajectory profiles. In: Hancock GR, Samuelsen KM, editors. *Advances in latent variable mixture models*. New York: Charlotte, NC: Information Age Publishing; 2008. p. 53–75.
- Finkel D, Andel R, Gatz M, Pedersen N. The role of occupational complexity in trajectories of cognitive aging before and after retirement. *Psychol Aging*. 2009;24:563–73. <https://doi.org/10.1037/a0015511>.
- Meyer J, Mutambudzi M. Construction of life-course occupational trajectories: evidence for work as a mediator of racial disparities in hypertension. *J Occup Environ Med*. 2012;54:1201–7. <https://doi.org/10.1097/JOM.0b013e31826bb6ac>.
- Magee C, Caputi P, Lee J. Distinct longitudinal patterns of absenteeism and their antecedents in full-time Australian employees. *J Occup Health Psychol*. 2016;21:24–36. <https://doi.org/10.1037/a0039138>.

16. Ubalde-Lopez M, Arends I, Almansa J, Delclos G, Gimeno D, Bültmann U. Beyond Return to Work: The Effect of Multimorbidity on Work Functioning Trajectories After Sick Leave due to Common Mental Disorders. *J Occup Rehabil*. 2016. [Epub ahead of print] <https://doi.org/10.1007/s10926-016-9647-0>.
17. von Bonsdorff ME, Kokko K, Seitsamo J, von Bonsdorff MB, Nygård C, Ilmarinen J, et al. Work strain in midlife and 28-year work ability trajectories. *Scand J Work Environ Health*. 2011;37:455–63. <https://doi.org/10.5271/sjweh.3177>.
18. Hou W, Sheu C, Liang H, Hsieh C, Lee Y, Chuang H, et al. Trajectories and predictors of return to work after traumatic limb injury—a 2-year follow-up study. *Scand J Work Environ Health*. 2012;38:456–66. <https://doi.org/10.5271/sjweh.3287>.
19. Laaksonen M, Metsä-Simola N, Martikainen P, Pietiläinen O, Rahkonen O, Gould R, et al. Trajectories of mental health before and after old-age and disability retirement: a register-based study on purchases of psychotropic drugs. *Scand J Work Environ Health*. 2012;38:409–17. <https://doi.org/10.5271/sjweh.3290>.
20. Wang M, Björkenstam C, Alexanderson K, Runeson B, Tinghög P, Mittendorfer-Rutz E. Trajectories of Work-Related Functional Impairment prior to Suicide. *PLoS One*. 2015;10:e0139937. <https://doi.org/10.1371/journal.pone.0139937>.
21. Amick BC, McLeod CB, Bültmann U. Labor markets and health: an integrated life course perspective. *Scand J Work Environ Health*. 2016;42(4):346–53. <https://doi.org/10.5271/sjweh.3567>.
22. López GM, Durán X, Zaballa E, Sanchez-Niubo A, Delclos G, Benavides F. Cohort profile: the Spanish WORKing life Social Security (WORKss) cohort study. *BMJ Open*. 2016;6:e008555. <https://doi.org/10.1136/bmjopen-2015-008555>.
23. Durán X, Martínez JM, Benavides FG. Occupational factors associated with the potential years of working life lost due to a non-work related permanent disability. *Work*. 2013;45:305–9.
24. Nylund KL, Asparouhov T, Muthén BO. Deciding on the Number of Classes in Latent Class Analysis and Growth Mixture Modeling: A Monte Carlo Simulation Study. *Structural Equation Modeling: A Multidisciplinary Journal*. 2007;14(4):535–69. <https://doi.org/10.1080/10705510701575396>.
25. Bauer DJ, Curran PJ. Overextraction of Latent Trajectory Classes: Much Ado About Nothing? Reply to Rindskopf (2003) Muthén (2003), and Cudeck and Henly (2003). *Psychological Methods*. 2003;8(3):384–93. <https://doi.org/10.1037/1082-989X.8.3.384>.
26. Muthén B. Statistical and substantive checking in growth mixture modeling: comment on Bauer and Curran (2003). *Psychol Methods*. 2003;8:369–77;384–93. <https://doi.org/10.1037/1082-989X.8.3.369>.
27. Rindskopf D. Mixture or Homogeneous? Comment on Bauer and Curran (2003). *Psychological Methods*. 2003;8(3):364–68. <https://doi.org/10.1037/1082-989X.8.3.364>.
28. Muthén B. Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data. In: Kaplan ID, editor. *Handbook of quantitative methodology for the social sciences*. Newbury Park, CA: Sage Publications; 2004. p. 345–68. <https://doi.org/10.4135/9781412986311.n19>.
29. Cebrían, I, Moreno, G. [The status of women in the Spanish labor market: mismatches and challenges] La situación de las mujeres en el Mercado de trabajo espa-ol: desajustes y retos. *Revista de Economía Industrial*. 2008;367:121–37. Spanish.
30. Wardenaar K, van LH, Cai T, Fava M, Gruber M, Li J, et al. The effects of co-morbidity in defining major depression subtypes associated with long-term course and severity. *Psychol Med*. 2014;44:3289–302. <https://doi.org/10.1017/S0033291714000993>.
31. Wardenaar K, Monden R, Conradi H, de Jonge P. Symptom-specific course trajectories and their determinants in primary care patients with Major Depressive Disorder: Evidence for two etiologically distinct prototypes. *J Affect Disord*. 2015;179:38–46. <https://doi.org/10.1016/j.jad.2015.03.029>.
32. López-Ruiz M, Artazcoz L, Martínez J, Rojas M, Benavides F. Informal employment and health status in Central America. *BMC Public Health*. 2015;15:698. <https://doi.org/10.1186/s12889-015-2030-9>.
33. Kristensen P, Aalen O. Understanding mechanisms: opening the black box in observational studies. *Scand J Work Environ Health*. 2013;39:121–4. <https://doi.org/10.5271/sjweh.3343>.

Received for publication: 14 July 2016