Gender inequalities at work in Southern Europe

Yijun Ren[1] · Alessandra Guglielmi[2] · Lara Maestripieri[3]

Abstract

Despite a long-term trend towards reduction, the gender gap in employment keeps standing in Southern Europe. Numerous potential causes have been individuated, such as the household configuration, women's human capital, or the institutions that regulate the labour market. Less is known about the role of the locality. This paper explores what covariates influence women's access to labour markets, and whether it is unevenly distributed across different countries and regions in Southern Europe. The analysis is based on the dataset round 9 (2018) from the European Social Survey. We focus on the following countries available in the dataset: Cyprus, Italy, Spain and Portugal. Italy and Spain are further differentiated into vulnerable and affluent regions according to the regional GDP in 2018. We apply a regression model for the binary response that is the indicator of having been doing paid work for the last 7 days of each individual in the sample. We adopt the Bayesian approach, to derive conclusions via a whole probability distribution, i.e., the posterior of all parameters, given data. The statistical goal is the selection of the most important covariates for access to the labour market, focusing on gender differences. Our analysis finds out that individual characteristics are mediated by household composition. Even though higher education increases women's employment, the presence of children and having an employed partner reduce such involvement. Moreover, a larger gender gap is detected in vulnerable regions rather than affluent ones, especially in Italy.

✉ Yijun Ren
  yijun.ren@polimi.it

  Alessandra Guglielmi
  alessandra.guglielmi@polimi.it

  Lara Maestripieri
  lara.maestripieri@gmail.com

[1] Department of Electronics, Information and Bioengineering, Politecnico di Milano, Milan, Italy

[2] Department of Mathematics, Politecnico di Milano, Milan, Italy

[3] Department of Political Science and Public Law/IGOP, Universitat Autónoma de Barcelona, Cerdanyola del Vallès, Spain

1 Introduction

Women's labour market access is an important societal, political and economic issue. In fact, despite the consistent steady growth that took place in European countries since the beginning of the '80s, contemporary labour markets are still characterised by gender inequalities in access to employment. However, women's disadvantage does not distribute evenly between countries, and, furthermore, even within countries there are strong territorial disparities that shape women's capacity to access employment. Southern Europe is the region that shows the persistence of the strongest inequalities in accessing employment between genders. If we look at the EIGE inequality index in the domain of access to work, we see that three out of four countries that traditionally are included in the Southern European model are well below the European average (81.3) and countries like France (83.7) or Germany (84.2): Spain (80.2), Greece (72.7) and Italy (69.1). The only exception to this rule is Portugal, with 88.2. The EIGE index[1] is a composite indicator proposed by the European Institute for Gender Equality, which allows a comparison between European countries in terms of gender equality. It includes several dimensions, although we have considered only the domain of access to work in this article. High values of this index denote more equal access to employment and good working conditions between men and women.

Women's access to the labour market is complicated by cultural reasons [1], but also by the specific configuration that the welfare state assumed in the Southern European countries [2]. In particular, the Southern European model of social protection [3] has always promoted a male breadwinner model, in which women's labour has been considered a secondary (and expendable) contribution to the main income that men procured to their household [4]. In this model, women are traditionally assigned to care and domestic role, while the state does not intervene to support the unpaid work provided by women [5]. Despite the unfavourable societal and political context that explains the lag in gender inequalities in Southern Europe, women's individual characteristics also take a role. In fact, the level of education, presence of children and family configuration affect the capacity of women to access the labour market [6]; the recent crises [7, 8] increased the inequality within women, as women most likely lose their jobs when they were the low educated, with a migrant background or in the presence of young children. In this paper, we aim to study what are the individual characteristics that might affect women's labour market participation and to what extent being a resident in one region or another can shape the relationship between gender and access to employment.

We consider data from the European Social Survey and focus on dataset round 9 (2018), the latest round available at the time of this research. This wave includes data among others on social conditions and indicators, inequality and social exclusion, elderly, youth, children, family life and marriage, with a special module on the timing of life, justice and fairness. The data come from interviews with respondents from Italy, Spain, Portugal and Cyprus, which are the only countries from South Europe in dataset round 9. We are interested in the probability that a respondent has been doing paid work in the last seven days (from the interview). The dataset collects further information, that we treat as covariates in a logistic regression model: gender, age, number of children, being an immigrant, education level, if the respondent has a partner, and if the partner has a job. We have also included a dummy variable for each country the respondents are from, splitting Spain and Italy into two parts.

We consider the Bayesian approach to the model since it automatically provided uncertainty quantification of the regression parameters. The posterior inference is computed through simulation methods easily obtained using probabilistic programming languages.

---

[1] https://eige.europa.eu/gender-equality-index/2021/compare-countries/work/1/bar.

The findings of the Bayesian logit model confirm that individual characteristics of women (e.g. education, migrant background) are mediated by household characteristics (e.g. having a partner, presence of children) in their influence on women's access to the labour market. Moreover, women's access to the labour market is unevenly distributed and depends on the different economic performances of regions.

The remainder of this paper is structured as follows. Section 2 describes all the difficulties for women to access the labour market in Southern Europe, while Sect. 3 introduces the dataset we have analyzed and a preliminary exploratory analysis. In Sect. 4 we present the Bayesian logit model we apply to the data. Section 5 presents the posterior inference and findings from this specific dataset. Section 6 concludes the paper with a discussion.

2 Women's labour market access in Southern Europe

2.1 Women's dilemma between paid and unpaid work

The role of family commitments in preventing women from accessing labour has always been debated from the feminist perspective on the gendered labour market. From one side, authors (as Williams [9]) argued that it is the role of the gendered unbalance in unpaid work (that is, care and domestic work) that prevented women from being active in the labour market as men do. Conversely, authors like Walby [10] sustained that the disadvantaged conditions of women's in the labour market exacerbated women's predominance in reproduction. The expression *Work-Family role system* indicates the interconnections of family, care policies and labour market regulation [11]; it indicates the strong degree of interdependence between the possibility that women have of taking up employment and the organisation of the family. If it is difficult to decide what comes first (gender roles or the labour market), exploring the relationships between the two is important.

This debate started already in the '70s, with the neoclassical theory of Becker [12] who sustained that the most rational choice for a household is to have one member specialised in paid activities and one member that specialised in unpaid activities necessary for reproduction. In his economic perspective, the most efficient equilibrium is given when the first were men and the second were women: men are usually paid more on labour markets, while women are *naturally* inclined to care, so they are more efficient in doing unpaid work. The fact that this household organization exacerbates gender inequality between men and women is not at all considered an issue in this theoretical perspective. A second important intervention was proposed by Hakim [13] with her theory of preferences. She focuses on analysing differences among women and puts forward two main ideal types of women confronting the *Work-Family role system*: the first group called *the grateful slaves* are women who prefer family over work commitment and the second one *the self-made women* have on the contrary a preference for work. She measures in a later work [14] that the two groups weigh 20% each, while about 60% of women have mixed preferences - that explains the variability in women's labour market access in real labour markets.

Hakim's intervention gave ground to fierce reactions, especially among feminist scholars. She was accused of blaming the victim [15], instead of pointing at the real constraints that women faced when trying to access the labour market: a gendered distribution of unpaid work, culturally grounded in the European societies [1] and a system of institutions that locked women into their role of mothers [16].

From a cultural perspective, gender norms prescribe the "correct" number of working hours for women: institutions are shaped over the gendered norms and reproduce inequalities that have their roots in how societies assign roles to genders [1]. From the institutional perspective, the reconciliation of work commitments outside of the home with family duties is strongly influenced by services available that effectively allow women to work. Women are thus structurally compelled to work less than men because labour markets are incompatible with unpaid work and the welfare state does not always offer adequate support for externalizing care outside families [15, 17].

Since family workloads are particularly significant when there are small children or elderly family members to look after, comparative studies reveal the importance of the availability of public childcare services and parental leave, if female employment is to be facilitated [5, 18, 19]. Many studies of the development of the welfare state from a gender perspective have shown, however, that the way in which childcare and elderly care becomes the responsibility of the State, varies enormously from one country to another [20]. In this regard, Naldini and Saraceno [21] utilise the concepts of *defamilisation* to describe the various degrees to which different welfare regimes pursue the redefinition of gender relations. To be more precise, the term *defamilisation* refers to the capillary presence of those services that release women in particular from the task of looking after dependant family members (e.g. young children), and thus increase their chances of pursuing a career which would otherwise be seriously threatened by the problem of reconciling work and family commitments. On the other hand, the term *supported familialism* means the existence of periods of leave, parental allowances and monetary transfers, which permit a person to interrupt work to look after a family member, thanks to the financial aid provided. In the absence of either type of policy, then *familialism by default* is the term used [5], which necessarily implies the family shouldering all responsibility for providing such care.

Geographically speaking, countries that are mora capable of reducing women's commitment to unpaid work are Scandinavian and Northern countries: in Scandinavian, the externalisation occurs thanks to the involvement of the State as a provider of services; in countries such as the UK or Ireland, it occurs through private service paid on the market. The supported familialism characterised Continental Europe, with countries like France, Germany or Austria extensively supporting families with monetary transfers and services in their reproductive role. Lastly, Southern European countries are characterised by the lack of support for families, with women entrapped in the *double shift* of being workers and mothers [22].

2.2 The specificity of Southern Europe

Maurizio Ferrera proposed the Southern European Welfare model in 1996. The original formulation of the welfare regime theory by Esping Andersen [23] included Southern Europe in the regime of Conservative countries. However, scholars like Ferrera were unsatisfied with this classification, as these countries were distinguished for several characteristics from the Conservative group. In fact, even if it is true that labour market policies are organised on a Bismarckian model of assurance like France or Germany [24], Southern Europe has a limited intervention in social care and family policies that completely distinguish these countries from Continental ones [22, 25]. Although following the social investment turn in European policies increased attention to the childcare service has also been taking place in the South, the provisions of childcare lag behind Continental and Northern Europe [26]. It

thus configures a situation in which there is a scarcity of publicly provided alternatives to family care, while the cash transfers to families with needy members are low.

The limited intervention of the state, defined by Chiara Saraceno the *unsupported familialism* [27], implies that the unpaid work is considered by the *Work-Family role system* as mainly a responsibility of families while the State intervene only when households are not able to comply their role. Consequently, countries in Southern Europe have limited provisions for care services, which jeopardises women's involvement in labour markets as they are usually the main provider of care work within families. In addition, the model of family that is assumed by the family policy is a model in which men are the primary earner, while women's income is a complement to men's one [4]. The male breadwinner model is persistent and quite far from the dual earner-dual carer model that lies behind the process of *defamilisation* [21].

As such, empirical evidence has demonstrated the importance of household composition in determining the likelihood of women's involvement in the labour market [7]. Together with education and the presence of children, having a partner becomes fundamental to predicting women's access to the labour market [28]. As childcare services are provided mostly at the local level, services are uneven between and within countries in Southern Europe. The strong territorial disparities in the provisions of services depend on the economic capacity of the regions and, even if the direction of causality is unclear, is strongly associated with women's access to the labour market [29]. Especially in bigger countries like Spain or Italy, geography shapes women's labour market access as much as individual characteristics, such as education, while regional differences are reduced in smaller countries such as Cyprus or Portugal.

In conclusion, based on the theory that we have reviewed regarding the *Work-Family role system* in Southern Europe we might elaborate the following hypothesis in relation to the determinants of women's access to the labour market:

I We expect that individual characteristics of women (e.g. education, migrant background) are mediated by household characteristics (e.g. having a partner, presence of children) in their influence on women's access to the labour market. Having a working partner and the presence of children reduce the likelihood that women work.

II Women's access to the labour market is unevenly distributed and depends on the different economic performances of regions. We assume that richer regions not only offer more employment opportunities but are also more capable of offering care services for working women.

3 Data description and exploratory analysis

In this section, we give a description of the data we focus on for later inference (in Sect. 5). In particular, Sect. 3.1 describes the dataset and variables used for analysis and Sect. 3.2 presents the exploratory data analysis which summarizes the main data characteristics by using statistical visualization tools.

3.1 Description of the dataset

The European Social Survey (ESS) is an academically driven cross-national survey that collects information in more than thirty European nations since its establishment in 2001. In

this paper we focus on dataset round 9 (2018) provided by the ESS; data can be downloaded at https://www.europeansocialsurvey.org/data/round-index.html.

As for similar rounds, the dataset results from an hour-long face-to-face interview for respondents aged 15 and over, who are residents within private households from twenty-seven European countries. The interview includes questions on various core topics: social trust, political interest and participation, socio-political orientation, social exclusion, etc. This results in a large number (557) of covariates for each respondent. Note that we will not use the whole dataset for analysis, since there are irrelevant covariates with respect to the main research question. We have organized the dataset as follows.

(a) We consider only those respondents aged between 25 and 55 years old since most people under 25 years old might be students and those over 55 might have retired yet or will be in a few years. This restriction is called *prime age*, and is adopted to focus on the most active stratum of population in the labour market. For instance, see [30].
(b) We restrict the analysis to data from countries of Southern Europe, where the gender gap in economic participation and opportunity is larger than in North Europe [31]. However not all South European countries participated in the survey. We have data from Cyprus, Italy, Portugal and Spain. This will be the definition of South Europe in the rest of the paper. However, we have excluded respondents from two Spanish regions, the enclaves of Ciudad Autonoma de Ceuta and Ciudad Autonoma de Melilla, as their economic situation is very different from other Spanish regions.

With the aim of improving readability, some variable labels we use in our model are different from those in the original European Social Survey dataset. We have also transformed some of the variables in the original dataset with the same aim. See "Appendix C". To sum up, we include the following variables in the dataset we focus on in this paper:

- *age*: respondent's age in years.
- *child*: number of children (son, daughter, step or adopted or foster child) under 14 years old of the respondent and living in the same household. We consider only children below 14 years old in this variable since it is well-known that they need more attention from parents, as they develop competencies, interests, a sense of confidence and autonomy during these years [32].
- *region*: region of the respondent's country.
- *country*: country of residence of the respondent. Instead of considering only the four countries in the dataset (Cyprus, Italy, Portugal, and Spain), we have considered two groups for Italy and Spain. Specifically, we have grouped Italian and Spanish regions into two levels, that is *affluent* and *vulnerable*, according to regional GDP per capita in 2018, since there exists a huge difference in terms of economic development between regions of these two countries [33]. More specifically, we calculated the median regional GDP per capita for Italy and Spain separately in 2018: the value equals 23,200 € for Spain and 31,700 € for Italy. The regions whose GDP per capita is lower than the median are considered *vulnerable* regions, otherwise *affluent* ones. See "Appendix A" for the associated classification. We use the covariate *country* in the rest of the paper, but the levels it assumes are *affluent* Italy, *vulnerable* Italy, *affluent* Spain, *vulnerable* Spain, Cyprus, and Portugal.
- *citizen*: respondent's citizenship status which can be autochthonous or immigrant. A respondent is autochthonous if he/she is born in the country and has citizenship of the country, otherwise, the person is classified as an immigrant. The variable *citizen* equals 1 if the respondent is autochthonous and 0 otherwise.

Table 1 Empirical mean and standard deviation of the response, *age* and *child*

Variable	Mean	Standard deviation
pdwrk	76%	0.009
age	41.55	8.77
child	0.57	0.82

- *education*: respondent's highest education level with three ordered levels. Education = 1 means up to lower secondary education or less; Education = 2 corresponds to upper secondary education or post-secondary or short-cycle tertiary; Education = 3 indicates other higher education levels.
- *gender*: respondent's gender.
- *partner*: status of respondent's partner, a categorical covariate with 4 levels. A respondent may live without a partner (*no partner*) or with an unemployed partner (*inactive partner*) or with a partner who looks actively for a job (*partner looking for job*) or with a partner who has been doing paid work for the last seven days (*working partner*).
- *pdwrk*: this variable is equal to 1 if the respondent performed at least one hour of paid work during the last seven days, 0 otherwise (definition by Eurostat[2]). This is the response variable in the model. In the rest of the paper, we denominate as *probability of employment* the probability that *pdwrk* is equal to 1.

The dataset downloaded from the ESS website is richer of information. First of all, according to the aim of the analysis, we selected only records with non-missing values of age and satisfying points (a) and (b), which resulted in a dataset of 2908 subjects. Based on experts' knowledge, we have initially included information on religion, sexual orientation and personal opinions, in addition to the covariates listed above. Then subjects with missing values among the larger set of selected covariates were further excluded; this resulted in a dataset from 2488 subjects. However, since too many non-informative covariates may influence the quality of the statistical inference we produce under our model, we have fitted preliminary frequentist lasso regression [34, 35] in a rather black-box fashion. This method automatically selects a subset of covariates that are most significant for the response. In our case, information on religion, sexual orientation and personal opinions were discarded as being non-significant.

The final dataset we consider for the Bayesian statistical analysis in this paper has the following characteristics: the sample size is $n = 2488$ (this is the number of respondents in the dataset), 8 independent variables (covariates, itemized above), and one dependent variable (*pdwrk*). The empirical sizes of the respondents in the four countries are $n_C = 293$ for Cyprus, $n_P = 437$ for Portugal, $n_{Iv} = 338$ for vulnerable Italian regions, $n_{Ia} = 478$ for affluent Italian regions, $n_{Sv} = 248$ for vulnerable Spanish regions and $n_{Sa} = 271$ for affluent Spanish regions, with $n_C + n_P + n_{Iv} + n_{Ia} + n_{Sv} + n_{Sa} = n$. We refer to this data as *SE (South Europe) labour market dataset*.

3.2 Exploratory data analysis

We start by reporting summary statistics for the response *pdwrk* and of the only numerical covariates, *age* and *child*. See Table 1.

---

[2] https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Employed_person-LFS.
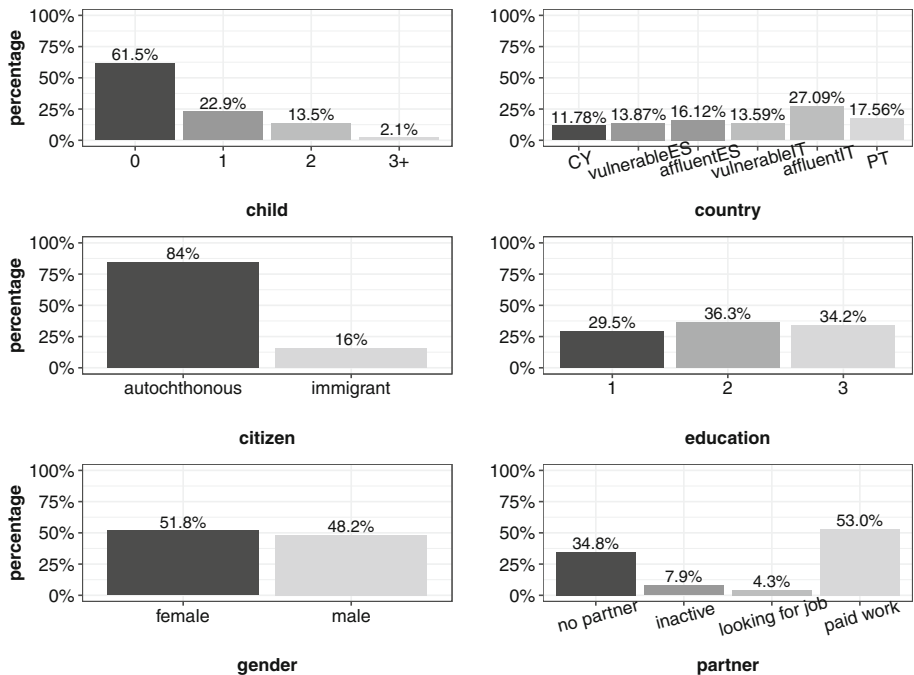
Fig. 1 Empirical frequencies of the categorical covariates in the dataset

Figure 1 shows the empirical frequencies of the categorical variables introduced in the previous section. Note that the numeric variable *child* has been discretized with levels *0, 1, 2, 3+*. Figure 1 shows that only 2.1% of respondents live with more than three children. As far as *citizen* is concerned, the dataset is clearly strongly unbalanced, since only 16% of respondents are immigrants. Other rounds of this data type, referring to years close to 2018, show the same unbalancedness. Differently from *citizen*, *gender* is balanced.

To build the model for the posterior analysis, we analyze potential relationships among the covariates. Figure 2 shows boxplots of *age* and *child* (the only numeric covariates in the sample) with respect to categorical covariates such as *citizen*, *education*, *partenr*, factorized by gender. From Fig. 2, we see that respondents with a high education level are younger, which might reflect the improvement of the education level in European countries during these last decades [36]. Moreover, the empirical distribution of *age* is shifted towards higher values for those respondents living with an inactive partner, especially for women; this could account for women with older partners who may have already retired. Respondents living without a partner and female respondents with an inactive partner have no child. Here we can observe a remarkable gender difference: the distribution of *child* is shifted toward higher values for men with an inactive partner than for women in the same condition.

To understand the dependency between categorical variables (except for *country* and *region*), we have performed chi-squared tests of independence; the associated *p* values are reported in Fig. 3. Of course, as a rule of thumb, the smaller the *p* value of the test in Fig. 3, the stronger the evidence against independence and the darker the colour we have used in the figure.
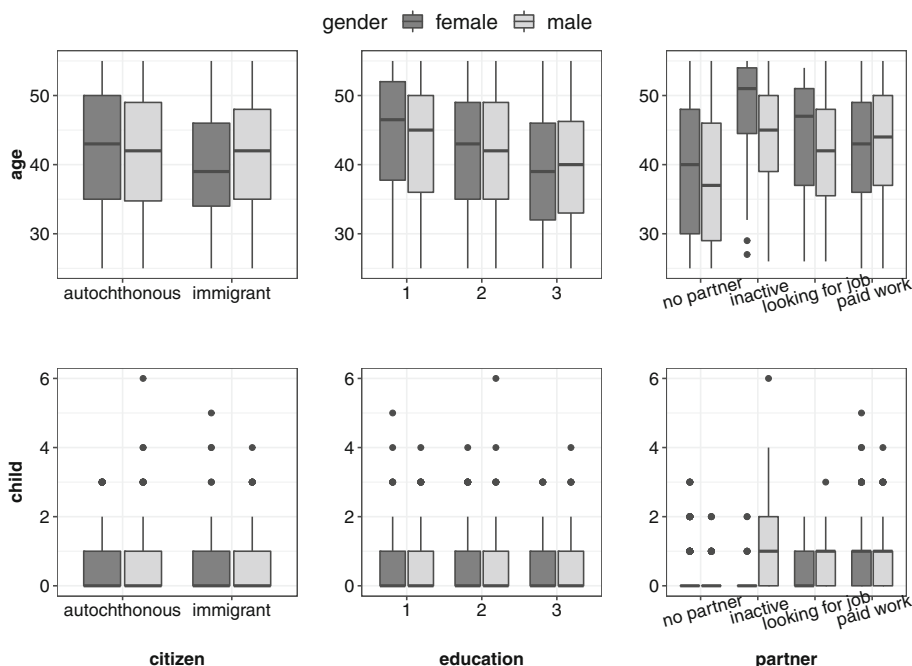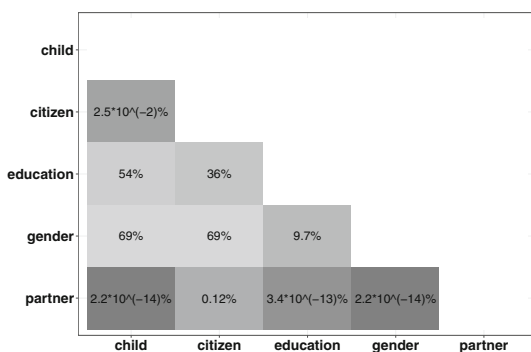
Fig. 2 Boxplots of *age* and *child* for each level of the categories of *citizen*, *education* and *partner*, factorized by gender

Fig. 3 *p* values of chi-squared tests between each pair of categorical variables. The smaller the *p* value, the stronger the evidence against independence and the darker the colour

It is clear that dependence is detected between *partner* on one side and each of *child* or *citizen* or *education* or *gender* on the other; see also the plots at the bottom row of Fig. 2. Similarly, there is empirical evidence of dependence between *child* and *citizen*. This is a well-known phenomenon, that immigrants tend to have a larger number of children [37].

Moreover, interactions between *age* and *education*, *age* and *partner* can be seen from Fig. 2.

4 Bayesian modelling of the impact of covariates

This section specifies the model we use in Sect. 5 for posterior inference on the SE labour market dataset, which is a Bayesian lasso regression model. We also include a brief description of the probabilistic programming language, Stan, that can build an appropriate Markov Chain Monte Carlo algorithm, to approximate the full posterior distribution.

Generalized linear models (GLM) were introduced by Nelder and Wedderburn [38] by extending ordinary regression models to non-normal response distributions. The logistic regression model, an example of GLMs, is widely used for modelling the relationship between the binary response variable and a set of explanatory variables. Since our response variables $Y_1, Y_2, \ldots, Y_n$ (an indicator of paid work for the survey respondents) are binary, we fit the following model with the logit link function:

$$Y_i | \pi_i \stackrel{\text{ind}}{\sim} Be(\pi_i), \quad i = 1, \ldots, n \tag{1}$$

where the logit link is given by

$$\text{logit}(\pi_i) = \boldsymbol{x}_i^t \boldsymbol{\beta}. \tag{2}$$

In (2), $\boldsymbol{x}_i = (1, x_{i1}, \ldots, x_{ip})$ is a $(p+1)$-dimensional vector of covariates for $i = 1, \ldots, n$, while $\boldsymbol{\beta}$ is a $(p+1)$-dimensional coefficient vector. The model is completed with the following prior distribution:

$$\beta_j | \lambda \stackrel{iid}{\sim} Laplace(0, 1/\lambda), \quad j = 0, \ldots, p \tag{3}$$

$$\lambda^2 \sim Gamma(a_\lambda, b_\lambda), \quad a_\lambda > 0, b_\lambda > 0. \tag{4}$$

This model represents a Bayesian counterpart of the lasso (least absolute shrinkage and selection operator) method introduced by Tibshirani [34]. The (frequentist) lasso method is a well-known regularization approach for generalized linear models. It shrinks some coefficients and sets others to 0, so automatically performs a subset selection by imposing a penalty term on the objective function. This penalty term is given by L_1-norm of the coefficient vector $\boldsymbol{\beta}$ multiplied by a parameter $\lambda \geq 0$, which expresses the amount of shrinkage. The larger λ, the larger the penalization on the magnitude of coefficients. The choice of $\boldsymbol{\beta}$ which minimizes the lasso penalty as described above corresponds to the Bayes posterior mode of $\boldsymbol{\beta}$ when the regression coefficients β_j's are a priori distributed as in (3). As suggested by Park and Casella [35], one can specify a diffuse hyperprior distribution for λ like in (4). This prior allows performing the variable selection automatically. This means that some coefficients get shrunk to exactly zero as they are not significant for the model. Another popular alternative for sparse Bayesian regression coefficient estimation is the horseshoe prior [39], characterized by its robustness at handling unknown sparsity and large outlying signals. Piironen and Vehtari [40] point out its two shortcomings—the absence of consensus on the specification of the global shrinkage hyperparameter and the problem of regularization for large parameters—and propose a new way to solve these problems. Moreover, the horseshoe prior has shown comparable performance in different circumstances with the popular two-component mixture priors known as the spike-and-slab [39].

For our application, we select the variables in Table 2 as covariate $\boldsymbol{x}_i$ in (2). They have been selected according to two criteria: (*i*) based on the preliminary analysis described in Sect. 3.2, and (*ii*) as those variables that are the most significant according to the full Bayesian models, i.e., including all variables and their two-way and three-way combinations. Specifically, this means that we have run our model (1)–(4) including all the covariates described in Sect. 3.1

Table 2 Independent variables and their combinations included in our statistical model

Covariate	Type	Covariate	Type
Age	Numeric	Child	Numeric
Citizen	Binary	Education	Categorical
Country	Categorical	Gender	Binary
Partner	Categorical		
Age·child	Numeric	Age·citizen	Numeric
Age·gender	Categorical	Child·education	Categorical
Child·gender	Categorical	Citizen·education	Categorical
Citizen·gender	Categorical	Education·gender	Categorical
Partner·age	Numeric	Partner·gender	Categorical
Age·child·education	Numeric	Age·child·gender	Numeric
Education·partner·age	Numeric	Education·partner·child	Numeric
Education·partner·gender	Categorical	Child·gender·education	Numeric

and all two-way and three-way combinations for few iterations, and have discarded all the two-way and three-way combinations that were clearly not significative. This results in all the combinations of variables listed in Table 2.

Note that all numerical variables have been standardized and categorical variables were transformed into dummies by using the Python function **dmatrix** from the **patsy** package [41]. For example, the covariate *gender* is equal to 1 for a female respondent and 0 otherwise. Hence the corresponding regression parameter expresses the algebraic increment on the logit of the success probability in case the respondent is female. The reference value, represented by the intercept term, corresponds to the logit of the success probability of a native man coming from Cyprus, with education level 1 and without a partner.

All the inference in the Bayesian approach is based on the posterior distribution, i.e., the conditional distribution of $(\boldsymbol{\beta}, \lambda)$, given data $y_1, \ldots, y_n$ and covariates $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$. In general, the posterior distribution is not available in closed analytic form; typical approximations are based on a class of simulation methods known as Markov Chain Monte Carlo (MCMC) [42]. However, MCMC methods can be extremely demanding from the computational point of view and the design of efficient MCMC algorithms is not a simple task. Thanks to the growth of computer processing power and the development of different accurate software products, and probabilistic programming languages, which can generate MCMC samples in a black-box fashion given the specification of a Bayesian model, Bayesian statistics has become more popular in recent years. Suitable probabilistic programming languages for MCMC methods are JAGS [43] and Stan [44, 45]. The default go-to software is Stan [46], which automatically generates samples from the target distribution, the posterior, using Hamiltonian Monte Carlo (HMC) and No-U-Turn Sampler (NUTS) that provide high-quality chain [47]. There is also an R interface, i.e. the package *rstan* [48], that we use to call the Stan sampler.

5 Posterior inference on the SE labour market dataset

In this section, we report posterior inference, such as posterior means and credibility intervals of parameters, for the model described in Sect. 4 and applied to our dataset presented in

Sect. 3.1. In this work, we fix $a_\lambda = 1$ and $b_\lambda = 0.1$ in (4); see Sect. 5.1 below. We have run a single chain by Stan for the model described in Sect. 4. The first 5,000 iterations have been discarded as burn-in and the last 5,000 have been saved for analysis. First, Sect. 5.2 reports the inference, while in Sect. 5.3 we report associated findings and comments.

5.1 Hyperparameters and simulation settings

We have fixed hyperparameter $a_\lambda = 1$ so that this implies that the prior for λ^2 is the exponential distribution, a rather flat prior, as asked in Park and Casella [35]. We computed standard goodness-of-fit indexes to fix hyperparameter b_λ. Specifically, we computed log-pseudo marginal likelihood (LPML) and widely applicable information criterion (WAIC) for our model under $b_\lambda = 0.01, 0.1, 1$. We found that the optimal value is $b_\lambda = 0.1$.

We have also fitted likelihood (1)-(2) with the horseshoe prior [39] by the R-package **rstanarm**; see Piironen and Vehtari (2017) for more detail. Posterior inference in terms of significant variables and associated credibility intervals are very similar to the posterior inference of our model as reported in the next section for the Bayesian lasso prior (3)–(4).

As mentioned, we report here the output from a single chain. The debate on one single longer chain versus multiple shorter chains is old. For instance, Geyer [49] support the former approach, because of computationally efficiency. Here we have examined trace and autocorrelation plots of all parameters, indicating that convergence was achieved.

5.2 Posterior means and 95% credible intervals

Figures 4 and 5 report posterior means and posterior 95% credibility intervals (CI) of each regression parameter. See also Tables 5 and 6 (in "Appendix B") which provide explicit numerical values of the posterior means and posterior 95% intervals of each parameter. As a selection criterion for covariates, we consider hard selection, that is if 0 is contained in the posterior 95% credibility intervals, then the corresponding covariate is not significative and can be removed from the model. We have signed in black all intervals which do not contain 0, while the rest are in light grey. However, to help the plots' readability, we have also signed in black those intervals that do not overlap the zero significantly, i.e. we consider the corresponding covariate to be significant enough. Figure 4 refers to the estimates of regression parameters corresponding to the linear terms and the two-way interactions, while Fig. 5 concerns estimates of parameters of three-way interactions.

From the left panel of Fig. 4, we see that covariates such as *child*, *region*, *country*, *education*, *gender* and *partner* are significant to explain the success probability, that is the probability of employment. Note that, since the posterior distributions of the regression parameters associated with *gender[female]* and *country[vulnerableIT]* are concentrated on negative values, being a woman coming from *vulnerable* Italian regions decreases the probability of employment. This confirms that there is a difference between regions within Spain and Italy: the probability of being employed in *vulnerable* Italian regions is smaller, but it is higher in *affluent* Italian and *affluent* Spanish regions, given the same value of all other covariates. Living with an inactive partner or having children less than 14 years old increases the probability of employment. Living with an employed partner increases the success probability as the posterior distribution of the regression coefficient of *partner[paid work]* is concentrated on the positive values. However, this is not true for women: the negative values assumed by the regression parameter assumed when we include the interaction between being a female respondent and having a working partner, i.e., partner[paid work]·gender[female],
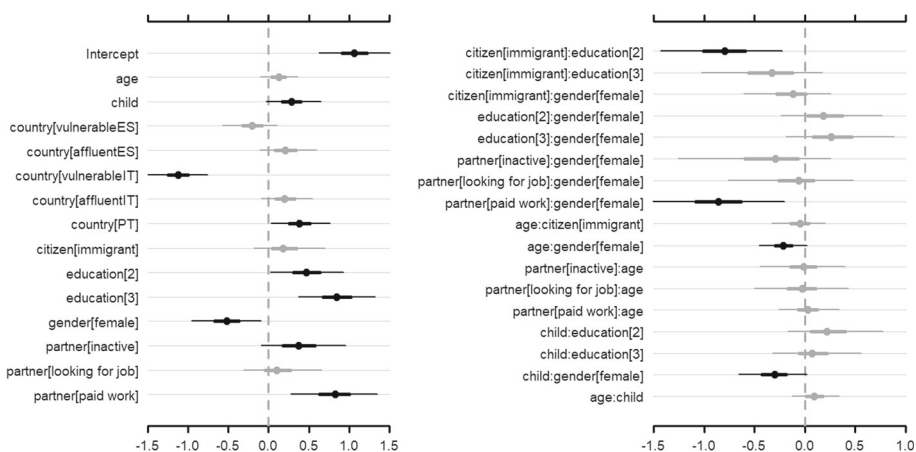
Fig. 4 Marginal posterior credible intervals for regression parameters; the points represent posterior means, the thicker segments are 50% CIs and the thinner ones are 95% CIs. Black intervals denote that the corresponding covariate is significative, i.e., the intervals do not contain 0
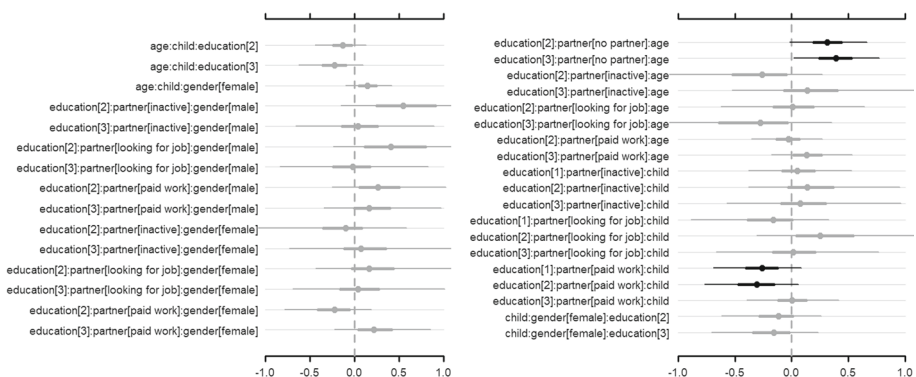

Fig. 5 Marginal posterior credible intervals for regression parameters; the points represent posterior means, the thicker segments are 50% CIs and the thinner ones are 95% CIs. Black intervals denote that the corresponding covariate is significative, i.e., the intervals do not contain 0

demonstrate the negative impact of having an employed partner on the employment probability for women; see the right panel of Fig. 4. The plot on the left of Fig. 4 shows that a higher education level increases the probability of being employed. Being an immigrant with a medium education level decreases the success probability as shown by the plot on the right, probably due to the problem of recognition of education obtained in other countries for immigrants.

From the right panel of Fig. 5, we observe other interesting interactions that involve the education level, such as education·partner·age and education·partner·child. A higher age can increase the probability of employment for respondents with education level 2 or 3 and without a partner. Having more children under 14 years old decreases the probability for respondents with education level 1 or 2 and living with an employed partner.

As a final remark in this section, we have fitted the likelihood (1)–(2) with the horseshoe prior using the R package **rstanarm**. Significative covariates do not change, and the credibility
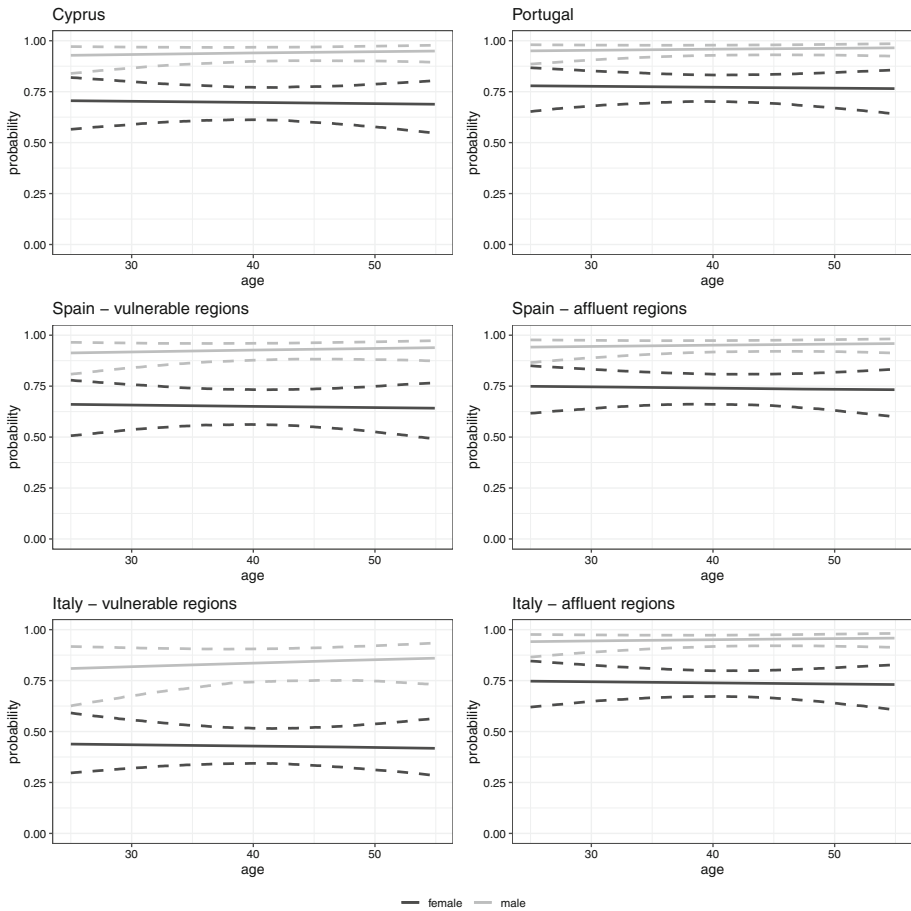
Fig. 6 Probability of been doing paid work in the last 7 days, as a function of *age* between 25 and 55 years old, with posterior 95% credible bands (dashed lines) for male and female reference respondents in different countries of Southern Europe

intervals of the regression parameters are very robust. For this reason, we do not report any further comments here.

5.3 Profiling respondents

In this section we focus on respondents with particular characteristics and calculate the associated probability of employment, to better understand how those significant covariates (described in Sect. 5.2) influence labour participation, possibly across different countries. The resulting plots do not show anything new regarding the inference displayed in Figs. 4 and 5, but they may help understand the gender gap in Southern European labour market.

First, let us consider as *reference* respondents two interviewed people with the following characteristics: not immigrant, having 1 child, with education level 2 and living with an employed partner. We denote these subjects as *female* and *male reference respondents*. Figure 6 shows the posterior medians (continuous lines) of the success probability for each
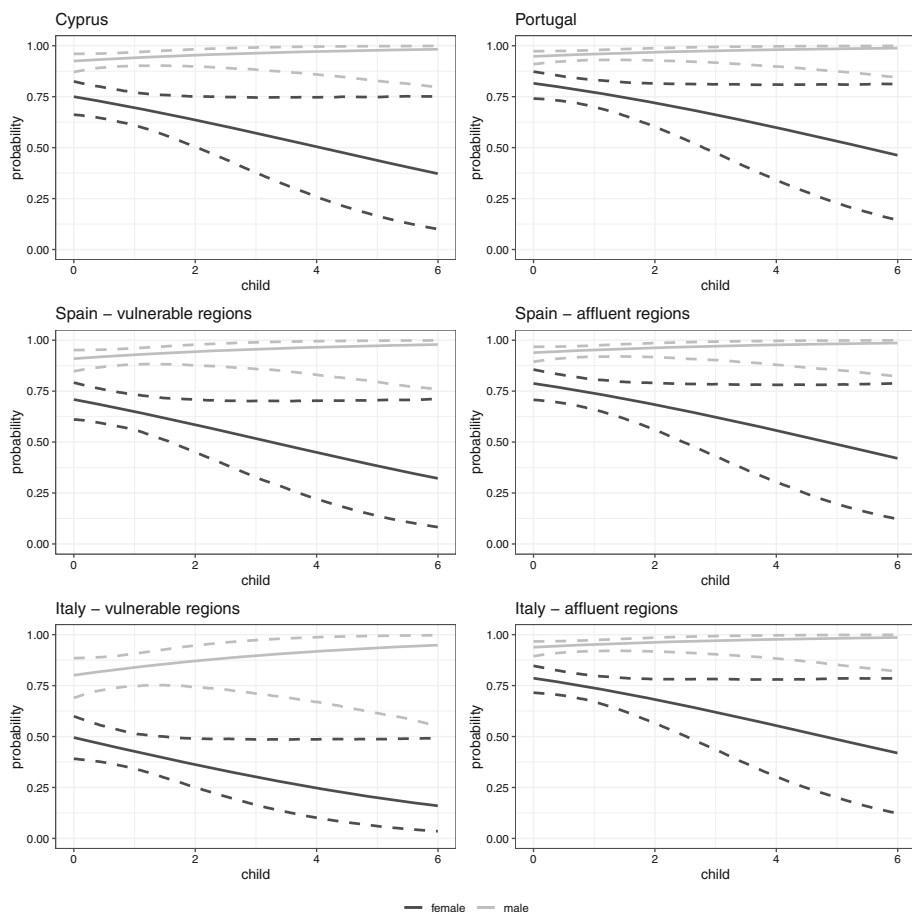
Fig. 7 Probability of been doing paid work for the last 7 days, as a function of *child* between 0 and 6, with posterior 95% credible bands (dashed lines) for male and female reference respondents in different countries of Southern Europe

value of *age* for the male and female reference respondents across different countries, while the dashed lines are the posterior 95% credibility bands. It is evident that there is a significant difference between genders: the success probability for female reference respondents is always lower than that for male respondents. A larger discrepancy can be observed in the vulnerable regions rather than the affluent ones.

Another significant covariate is *child* and we show the success probabilities, as a function of this covariate, for two *reference* respondents with the following characteristics: not immigrant, 42 years old (empirical median of *age*), with education level 2 and living with an employed partner. Figure 7 shows the success probability with 95% posterior credible bands in dashed lines for the female and male reference respondents as a function of the number of children. As a result, as the number of children increases, women are more excluded by the labour market, especially in *vulnerable* Italian regions, while men are more likely to be employed. The gender differences become more evident with a higher number of children.

From Fig. 4 (right) we have seen that one of the significative two-way combinations is the interaction between *citizen* and *education*. Table 3 shows the posterior mean of the employment probability for a female/male reference respondent with different education levels, citizenship and countries, given that both are 42 years old, have 1 child and live with an employed partner. If we condition respondents coming from the same country, with the same citizenship status and the same education level, the success probability for any male reference respondent is always higher than that for the associated female respondent. While for the male reference respondents it is true that the higher the education level, the higher the probability of being employed, regardless of being *immigrant* or not, this is not true for female reference respondents. In fact, in *vulnerable* Italian or Spanish regions, female immigrants with education level 2 are more likely to be excluded from the labour market than those with education level 1. Moreover, the gender difference is particularly evident in *vulnerable* Italian regions since the employment probability for female respondents is much lower than that for male respondents.

We now focus on significative three-way interactions pointed out by Fig. 5 such as the association between *education*, *partner* and *age* and between *education*, *partner* and *child*. To visualize how the first combination affects the probability of being employed, we have computed such probability as a function of *age* for reference respondents who are not immigrants, have 1 child and live in *vulnerable* Italian regions. The plots are reported in Fig. 8; in this case, the 95% posterior credible bands were not included in the figure to make the plots clearer. From Fig. 8, it is evident that a higher education level increases the employment probability. In particular, education can bring radical changes to women living with an employed partner: female reference respondents with an education level equal to 1 have always a lower probability of being employed than others in the same condition but with different partner's status (the plot in the top right corner). However, if they are highly educated, i.e., with education level 3, they are more active in the labour market than those living without a partner or with an inactive partner (the plot in the bottom right corner). A larger gender discrepancy can be observed by comparing the success probability for male and female reference respondents with education level 1 (the plots on top): while male respondents with an employed partner are more active in the labour market with respect to other male respondents with different partner's status, female respondents with an employed partner have lower success probability than others. We also observe that ageing is a helpful factor for single respondents, in particular for women with high education level: while the curve of the employment probability for single female reference respondents with education level 1 is almost flat as age varies, it is clearly increasing for women with education levels 2 or 3.

To display the interaction between *education*, *partner* and *child*, Fig. 9 reports the employment probability curves as a function of *child* for male and female respondents with any education level and any *p*artner's status from *vulneerable* Italian regions (not immigrant, 42 years old). While the trend of the curves remains similar for male reference respondents with different education levels or partner statuses, it changes significantly for female reference respondents. The probability of being employed decreases for women living with an employed partner as the number of children increases, especially if women have a low education level. There exists a great difference between female reference respondents with education level 2, but characterized by different partner's status: in case their partners are inactive or looking for a job, they tend to become more active in the labour market with a higher number of children and their probability curves share similar tendencies with those of male reference respondents with education level 2, while they are more excluded by the labour market if living with an employed partner. Note also that the employment probability curves for women living with an employed partner and with a low education level (level 1 or

Table 3 Posterior estimates of success probability for female and male reference respondents varying citizenship and education level

| Country	Citizen	Education	Gender	
			Female	Male
Cyprus	Autochthonous	1	0.59	0.86
		2	0.69	0.94
		3	0.86	0.96
	Immigrant	1	0.64	0.89
		2	0.56	0.89
		3	0.84	0.95
Spain—vulnerable regions	Autochthonous	1	0.54	0.84
		2	0.65	0.93
		3	0.83	0.95
	Immigrant	1	0.59	0.87
		2	0.51	0.87
		3	0.81	0.94
Spain—affluent regions	Autochthonous	1	0.64	0.89
		2	0.74	0.95
		3	0.88	0.97
	Immigrant	1	0.68	0.91
		2	0.61	0.91
		3	0.87	0.96
Italy—vulnerable regions	Autochthonous	1	0.32	0.69
		2	0.43	0.84
		3	0.67	0.88
	Immigrant	1	0.37	0.73
		2	0.29	0.74
		3	0.63	0.86
Italy—affluent regions	Autochthonous	1	0.64	0.89
		2	0.74	0.95
		3	0.88	0.97
	Immigrant	1	0.68	0.91
		2	0.61	0.91
		3	0.87	0.96
Portugal	Autochthonous	1	0.68	0.91
		2	0.77	0.96
		3	0.90	0.97
	Immigrant	1	0.72	0.92
		2	0.65	0.93
		3	0.88	0.97
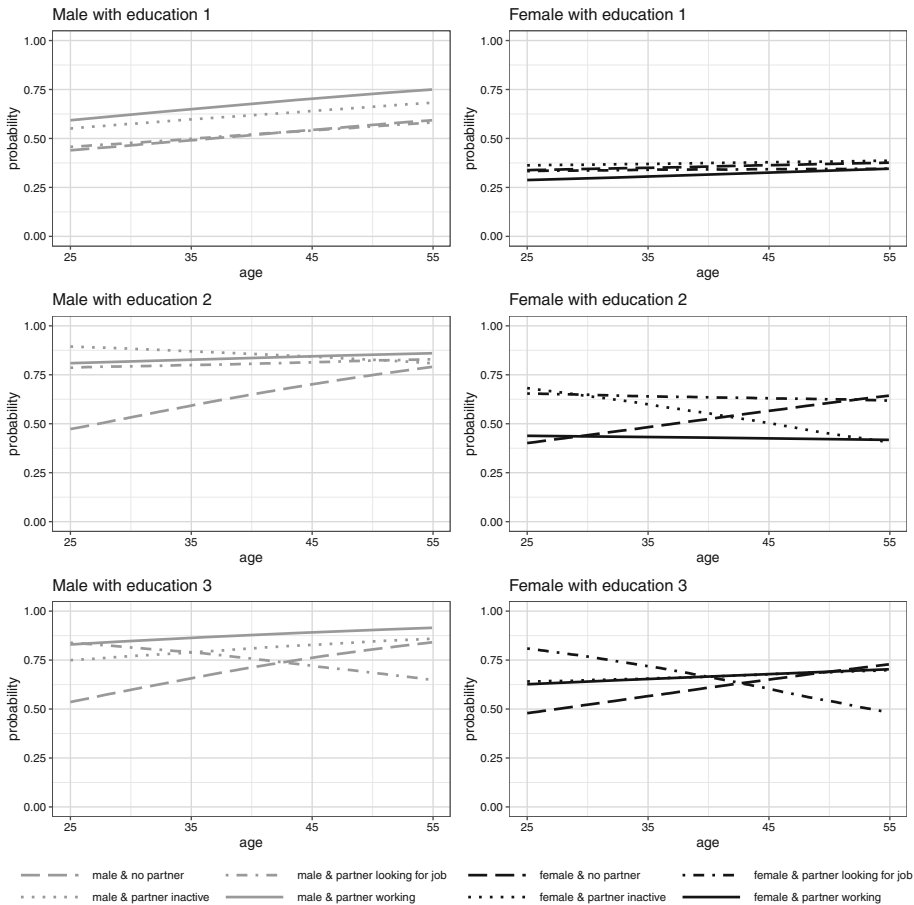
Fig. 8 Probability of been doing paid work in the last 7 days for male and female reference respondents varying age, partner's status and education level in *vulnerable* Italian regions

2) are quite similar, but they differ significantly from the curve for women highly educated, i.e., with education level 3.

We have repeated the same analysis for reference respondents living in other countries. However, the gender difference is more evident for respondents living in *vulnerable* Italian regions and this is the reason why we have focused on this case here.

6 Discussion and future work

In this article, we have analyzed the determinants of women's access to the labour market in South Europe, focusing on Cyprus, Italy, Spain and Portugal. Our findings confirm that individual characteristics of women are mediated by household characteristics in their influence on women's involvement in the labour market. On the one hand, education significantly impacts access to the labour market for both genders, especially for women; the gender gap narrows among highly educated people. Whilst ageing can be seen as a helpful factor for men
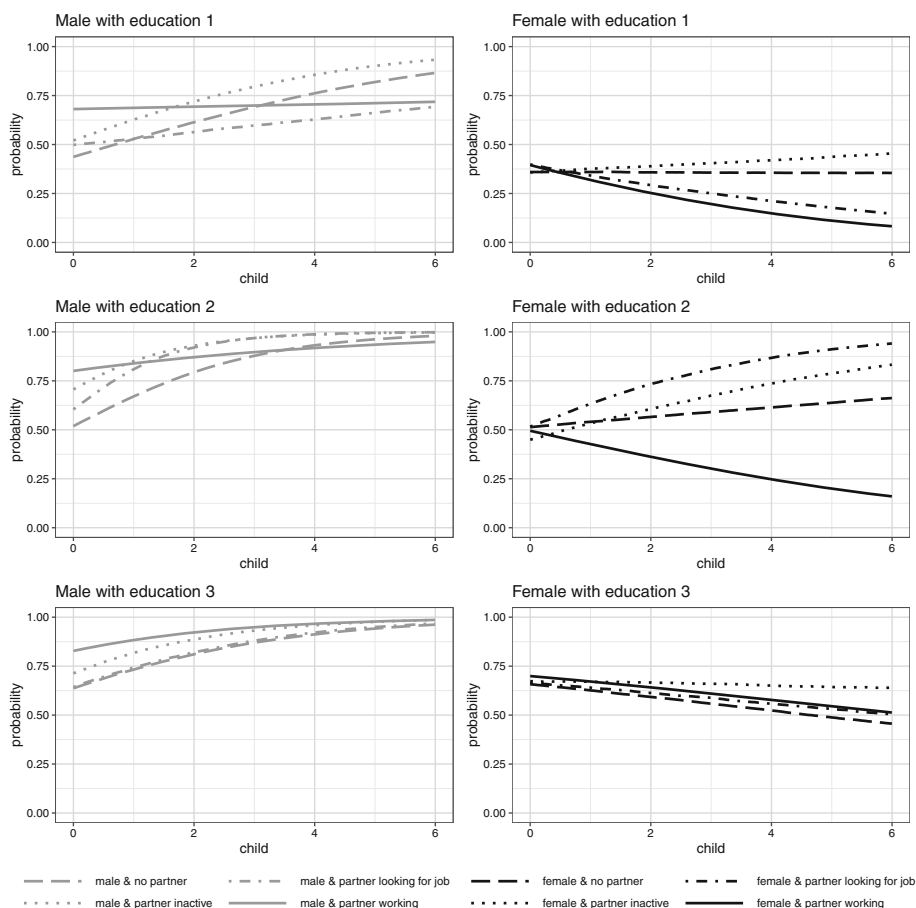
Fig. 9 Curve of the employment probability for male and female reference respondents varying number of children, partner's status and education level in *vulnerable* Italian regions

being employed, this does not hold for women. On the other hand, the presence of children and having a partner, particularly having an employed partner, reduce the probability that women work.

The situation becomes more delicate and complicated when taking into account combinations of individual and household characteristics. The gender gap in terms of the access to labour market becomes negligible among highly educated people without partners. For these subjects, the probability of employment increases with ageing, regardless of gender. The presence of children has a significant negative impact on access to the labour market for women with low or medium education levels and having an employed partner. In these families, women usually provide family-related work, while men take the role of the breadwinner due to the limited provision of care services in Southern Europe.

Moreover, our findings confirm the uneven distribution of labour market access in South Europe, in accordance with the EIGE inequality index in the domain of access to work: among the Southern European countries included in our dataset, Portugal outperforms other countries in terms of access to the labour market. This uneven distribution is strongly dependent on the

economic capacity of regions or countries; there are sharp differences between vulnerable and affluent regions, especially in Italy. Moreover, our analysis points out the persistence of gender inequality in South Europe, more evident in vulnerable Italian regions rather than in other regions. This can be traced back to the radical regional difference in Italy. An attempt to reduce the highlighted differences in the supply of early childcare services was made by the nationwide program, "Piano Straordinario per lo Sviluppo dei Servizi per la Prima Infanzia". However, the program effectiveness was scarce in the Southern regions, and stronger in other regions [50].

We also note that subjects with an immigrant background, and a low education level are more likely to be employed than native-born subjects with low education. But the situation is reversed in case subjects received middle or high education. Indeed, the competencies acquired by immigrants in their country of origin may be not suitable for the labour market of the current country due to cultural or language barriers. Thus, access to highly qualified occupations for immigrants might be discouraged by such differences, while access to occupations that do not require specific skills, i.e. manual workers, is more straightforward. Moreover, the effect of education is not linear among immigrants: subjects with medium education levels are less active in the labour market than those with a low education level. Here one should consider the problem of recognition of education obtained in other countries and the mismatch between jobs and skills among immigrants [37].

In this work, we have assumed that the effect of being from a country (including the classification of Spain and Italy into two groups each) is fixed. In future work, we could include this as random effects in a generalised linear mixed-effects model, thus assuming a hierarchical model where groups might exchange a priori information as it is customary in these types of Bayesian models. See [51] for an approach of this type on a similar dataset.

Declarations

Appendix A List of affluent and vulnerable regions of Italy and Spain

As mentioned in Sect. 3.1, we have calculated the median regional GDP per capita for Italy and Spain separately in 2018: the value equals 23,200 € for Spain and 31,700 € for Italy. The regions whose GDP per capita is lower than the median are considered *vulnerable* regions, otherwise *affluent* ones. This classification procedure results in Table 4.

Table 4 List of regions classified as *affluent* or *vunerable* regions in Italy and Spain

Country	Regions
Italy—affluent regions	Friuli Venezia Giulia, Emilia Romagna, Lazio, Liguria, Lombardia, Marche, Piemonte, Toscana, Trentino Alto Adige, Umbria, Veneto, Valle d'Aosta
Italy—vulnerable regions	Abruzzo, Basilicata, Calabria, Campania, Molise, Puglia, Sardegna, Sicilia
Spain—affluent regions	Aragón, Castilla y León, Cataluña, Cantabria, Comunidad de Madrid, Comunidad Foral de Navarra, Illes Baleares, Galicia, La Rioja, País Vasco
Spain—vulnerable regions	Andalucía, Canarias, Catilla-La Mancha, Comunidad Valenciana, Extremadura, Principado de Asturias, Región de Murcia

Appendix B Posterior means and 95% credibility intervals

Tables 5 and 6 report posterior means and marginal posterior 95% credible intervals for all regression parameters in model (1)–(4).

Appendix C Variable elaboration

As mentioned in Sect. 3.1, some variable labels we use in this manuscript are different from those in the original dataset from the European Social Survey, dataset round 9 (2018). To help reproducibility, we report the match in Table 7. We have also used variables as combinations of two or more variables from the original dataset. We explain here below how we have built them:

- *child*: the original dataset does not explicitly contain a variable that describes the number of children of the respondent, but it provides the description of the relationship between the respondent and everyone else living in the same household through the categorical variables rshipa2,..., rshipa15. Their age is registered by the numeric variables yrbrn2,..., yrbrn15. The number of children under 14 years old of respondents can be inferred from this information resulting in a new numeric variable *child*.
- *country*: this variable is derived from the categorical variable *region* and created according to Table 4.
- *citizen*: this variable indicates the citizenship of respondents and is created by combining the independent binary variables *ctzcntr* (citizen of the country?) and *brncntr* (born in the country?). The respondents are divided into two groups, autochthonous and immigrant, as specified in Table 8. Hence this variable is binary.
- *education*: the categorical variable *edulvlb* in the original dataset from the European Social Survey represents the education level of respondents according to ISCED classification and contains a 3-digit hierarchical coding framework that summarizes different education systems of all participating countries. The first digit of the code represents the 8ISCED11 levels, while the second and third digits specify educational programmes within levels. In order to simplify this variable, a new categorical variable *education* has been created focusing on the first digit of *edulvlb* and described in Table 9.
- *partner*: this variable has been introduced to better specify the partner's status which may influence significantly the chance of participating in the labour market, especially for women. So from the binary variables *pdwrkp, edctnp, uemplap, uemplip, dsbldp, rtrdp, cmsrvp* and *hswrkp* in the original dataset, we are able to derive the partner's

Table 5 Posterior means and 95% credible intervals of β

Covariate	Posterior	95% credible intervals	
	Mean	Lower bound	Upper bound
Intercept	1.07	0.63	1.51
Age	0.13	− 0.1	0.37
Child	0.29	− 0.03	0.66
Country [vulnerableES]	− 0.21	− 0.57	0.11
Country [affluentES]	0.22	− 0.11	0.6
Country [vulnerableIT]	− 1.12	− 1.5	− 0.76
Country [affluentIT]	0.21	− 0.09	0.55
Country [PT]	0.39	0.03	0.77
Citizen [immigrant]	0.21	− 0.18	0.71
Education [2]	0.47	0.02	0.94
Education [3]	0.85	0.37	1.32
Gender [female]	− 0.52	− 0.96	− 0.1
Partner [inactive]	0.39	− 0.1	0.96
Partner [looking for job]	0.12	− 0.31	0.66
Partner [paid work]	0.82	0.27	1.35
Citizen [immigrant]·education [2]	− 0.8	− 1.43	− 0.22
Citizen [immigrant]·education [3]	− 0.36	− 1.03	0.17
Citizen [immigrant]·gender [female]	− 0.14	− 0.61	0.26
Education [2]·gender [female]	0.21	− 0.24	0.77
Education [3]·gender [female]	0.29	− 0.19	0.89
Partner [inactive]·gender [female]	− 0.35	− 1.25	0.26
Partner [looking for job]·gender [female]	− 0.09	− 0.76	0.48
Partner [paid work]·gender [female]	− 0.86	− 1.5	− 0.2
Age·citizen [immigrant]	− 0.05	− 0.32	0.21
Age·gender [female]	− 0.21	− 0.45	0.02
Partner [inactive]·age	− 0.02	− 0.44	0.4
Partner [looking for job]·age	− 0.03	− 0.5	0.44
Partner [paid work]·age	0.03	− 0.25	0.34
Child·education [2]	0.24	− 0.17	0.78
Child·education [3]	0.09	− 0.32	0.56
Child·gender [female]	− 0.3	− 0.65	0.02
Age·child	0.1	− 0.12	0.34

main activity in the last seven days. Then, based on this information and on *prtnr* which indicates whether the respondent has a partner, the new variable *partner* has been defined as in Table 10.

Table 6 Posterior means and 95% credible intervals of β

Covariate	Posterior Mean	95% credible intervals	
		Lower bound	Upper bound
Age·child·education [2]	− 0.14	− 0.44	0.13
Age·child·education [3]	− 0.23	− 0.62	0.1
Age·child·gender [female]	0.15	− 0.1	0.42
Education [2]·partner [inactive]·gender [male]	0.61	− 0.15	1.66
education [3]·partner [inactive]·gender [male]	0.07	− 0.66	0.89
Education [2]·partner [looking for job]·gender [male]	0.5	− 0.24	1.73
Education [3]·partner [looking for job]·gender [male]	− 0.03	− 0.83	0.83
Education [2]·partner [paid work]·gender [male]	0.3	− 0.25	1.02
Education [3]·partner [paid work]·gender [male]	0.21	− 0.34	0.97
Education [2]·partner [inactive]·gender [female]	− 0.15	− 1.09	0.58
Education [3]·partner [inactive]·gender [female]	0.14	− 0.73	1.33
Education [2]·partner [looking for job]·gender [female]	0.22	− 0.44	1.09
Education [3]·partner [looking for job]·gender [female]	0.07	− 0.69	1.01
Education [2]·partner [paid work]·gender [female]	− 0.25	− 0.78	0.19
Education [3]·partner [paid work]·gender [female]	0.25	− 0.23	0.85
Education [2]·partner [no partner]·age	0.32	− 0.02	0.67
Education [3]·partner [no partner]·age	0.39	0.02	0.77
Education [2]·partner [inactive]·age	− 0.3	− 1.09	0.27
Education [3]·partner [inactive]·age	0.19	− 0.53	1.14
Education [2]·partner [looking for job]·age	0.01	− 0.62	0.64
Education [3]·partner [looking for job]·age	− 0.37	− 1.49	0.35
Education [2]·partner [paid work]·age	− 0.03	− 0.35	0.27
Education [3]·partner [paid work]·age	0.15	− 0.18	0.53
Education [1]·partner [inactive]·child	0.06	− 0.38	0.53
Education [2]·partner [inactive]·child	0.19	− 0.38	0.96
Education (3]·partner [inactive]·child	0.12	− 0.57	0.96
Education [1]·partner [looking for job]·child	− 0.2	− 0.89	0.33
Education [2]·partner [looking for job]·child	0.32	− 0.31	1.28
Education [3]·partner [looking for job]·child	0.03	− 0.67	0.77
Education [1]·partner [paid work]·child	− 0.27	− 0.69	0.08
Education [2]·partner [paid work]·child	− 0.32	− 0.77	0.06
Education [3]·partner [paid work]·child	0.01	− 0.39	0.41
Child·gender [female]·education [2]	− 0.14	− 0.62	0.26
Child·gender [female]·education [3]	− 0.19	− 0.71	0.23

Table 7 Match between the original variable labels in the original dataset from the European Social Survey, dataset round 9 (2018) and those in our model

Original label	Modified label
Agea	Age
Region	Region
gndr	Gender
Pdwrk	pdwrk

Table 8 Definition of the variable *citizen*

Citizen	brncntr	ctzcntr
Autochthonous	Yes/no	Yes
	Yes	Yes/no
Immigrant	No	No

Table 9 Definition of the variable *education*

Education	First digit of edulvlb	Education level
1	0	Less than primary education
	1	Primary education
	2	Lower secondary education
2	3	Upper secondary education
	4	Post-secondary non tertiary
	5	Short-cycle tertiary
3	6	Bachelor or equivalent
	7	Master or equivalent
	8	Doctoral or equivalent

Table 10 Definition of the variable *partner*

Partner	Condition	Partner's main activity for the last 7 days
No partner	prtnr = 0	Respondent lives without a partner
Inactive	edctnp=1	Partner has been doing education
	uemplip = 1	Partner has been unemployed and not looking for jobs
	dsbldp = 1	Partner has been permanently sick or disabled
	rtrdp = 1	Partner has been retired
	hswrkp = 1	Partner has been doing household chores
Looking for job	uemplap = 1	Partner has been unemployed and looking for jobs
Paid work	pdwrkp = 1	Partner has been doing paid work
	cmsrvp = 1	Partner has been doing community/military services

References

1. Pfau-Effinger, B.: Gender cultures and the gender arrangement-a theoretical framework for cross-national gender research. Innov. Eur. J. Soc. Sci. Res. **11**(2), 147–166 (1998)
2. Guillén, A.M., Jessoula, M., Matsaganis, M., Branco, R., Pavolini, E.: Southern European welfare systems in transition. In: Mediterranean Capitalism Revisited. One Model, Diferent Trajectories, 149 (2022)
3. Ferrera, M.: The southern model of welfare in social Europe. J. Eur. Soc. Policy **6**(1), 17–37 (1996). https://doi.org/10.1177/095892879600600102
4. Vesan, P.: Ancora al sud? I paesi mediterranei e le riforme delle politiche del lavoro negli anni della crisi economica. Meridiana **83**, 91–119 (2015)
5. Saraceno, C., Keck, W.: Towards an integrated approach for the analysis of gender equity in policies supporting paid work and care responsibilities. Demogr. Res. **25**, 371–406 (2011). https://doi.org/10.4054/DemRes.2011.25.11
6. Akkan, B.: An egalitarian politics of care: young female carers and the intersectional inequalities of gender, class and age. Fem. Theory **21**(1), 47–64 (2020). https://doi.org/10.1177/1464700119850025

7. Sánchez-Mira, N., O'Reilly, J.: Household employment and the crisis in Europe. Work Employ Soc. **33**(3), 422–443 (2019). https://doi.org/10.1177/0950017018809324
8. Maestripieri, L.: The covid-19 pandemics: why intersectionality matters. Front. Sociol. **6**, 642–662 (2021). https://doi.org/10.3389/fsoc.2021.642662
9. Williams, J.: Unbending Gender: Why Family and Work Conflict and What to do About It. Oxford University Press, Oxford (2001)
10. Walby, S.: Patriarchy at Work Minneapolis. University of Minnesota Press, Minneapolis (1986)
11. Pleck, J.H.: The work-family role system. Soc. Probl. **24**(4), 417–427 (1977). https://doi.org/10.2307/800135
12. Becker, G.S.: A theory of marriage: part I. J. Polit. Econ. **81**(4), 813–846 (1973)
13. Hakim, C.: Grateful slaves and self-made women: fact and fantasy in women's work orientations. Eur. Sociol. Rev. **7**(2), 101–121 (1991). https://doi.org/10.1093/oxfordjournals.esr.a036590
14. Hakim, C.: Work-Lifestyle Choices in the 21st Century: Preference Theory. Oxford University Press, Oxford (2000)
15. Ginn, J., Arber, S., Brannen, J., Dale, A., Dex, S., Elias, P., Moss, P., Pahl, J., Roberts, C., Rubery, J.: Feminist fallacies: a reply to hakim on women's employment. Br. J. Sociol. **47**(1), 167–174 (1996)
16. Fagan, C., Rubery, J.: The salience of the part-time divide in the European Union. Eur. Sociol. Rev. **12**(3), 227–250 (1996)
17. McRae, S.: Constraints and choices in mothers' employment careers: a consideration of Hakim's preference theory. Br. J. Sociol. **54**(3), 317–338 (2003). https://doi.org/10.1080/0007131032000111848
18. Del Boca, D., Pasqua, S., Pronzato, C.: Motherhood and market work decisions in institutional context: a European perspective. Oxf. Econ. Pap. **61**(suppl_1), 147–171 (2008)
19. Plantenga, J., Remery, C.: Reconciliation of work and private life. In: Gender and the European Labour Market, pp. 108–123. Routledge, Abingdon, Oxon (2013)
20. Lewis, J.: Gender and the development of welfare regimes. J. Eur. Soc. Policy **2**(3), 159–173 (1992). https://doi.org/10.1177/095892879200200301
21. Naldini, M., Saraceno, C.: Conciliazione famiglia e lavoro. Il Mulino, Bologna (2011)
22. Naldini, M.: The Family in the Mediterranean Welfare States, 1st edn. Routledge, London (2003). https://doi.org/10.4324/9780203009468
23. Esping-Andersen, G.: The Three Worlds of Welfare Capitalism. Polity Press, Cambridge (1990)
24. Palier, B.: Continental Western Europe—the Bismarckian welfare systems. In: Leibfried, S., Lewis, J., Obinger, H., Pierson, C. (eds.) The Oxford Handbook of the Welfare State, pp. 601–615. Oxford University Press, Oxford (2010)
25. Guillén, A.M., León, M.: The Spanish Welfare State in European Context. Ashgate Publishing Ltd, Burlington (2011)
26. León, M., Pavolini, E.: 'Social investment' or back to familism: the impact of the economic crisis on family and care policies in Italy and Spain. South Eur. Soc. Polit. **19**(3), 353–369 (2014). https://doi.org/10.1080/13608746.2014.948603
27. Saraceno, C.: Varieties of familialism: comparing four Southern European and East Asian welfare regimes. J. Eur. Soc. Policy **26**(4), 314–326 (2016). https://doi.org/10.1177/0958928716657275
28. Kasearu, K., Maestripieri, L., Ranci, C.: Women at risk: the impact of labour-market participation, education and household structure on the economic vulnerability of women through Europe. Eur. Soc. **19**(2), 202–221 (2017). https://doi.org/10.1080/14616696.2016.1268703
29. Arlotti, M., Sabatinelli, S.: Care as multi-scalar policy: Ecec and ltc services across Europe. In: Handbook on Urban Social Policies, pp. 117–133. Edward Elgar Publishing, Cheltenham (2022)
30. Huete-Morales, M.D., Vargas-Jiménez, M.: Modelling part-time employment in Spain: do women opt for fewer hours or do they have no choice? J. Gend. Stud. **27**(7), 815–833 (2018)
31. Schwab, K., Samans, R., Zahidi, S., Leopold, T.A., Ratcheva, V., Hausmann, R., Tyson, L.D.: The Global Gender Gap Report 2017, pp. 3–14. World Economic Forum, Cham (2017)
32. Eccles, J.S.: The development of children ages 6–14. Future Child. **9**, 30–44 (1999)
33. Eurostat: Regional GDP per Capita Ranged from 30% to 263% of the EU Average in 2018. https://ec.europa.eu/eurostat/web/products-euro-indicators/-/1-05032020-ap
34. Tibshirani, R.: Regression shrinkage and selection via the lasso. J. R. Stat. Soc. Ser. B (Methodol.) **58**(1), 267–288 (1996)
35. Park, T., Casella, G.: The Bayesian lasso. J. Am. Stat. Assoc. **103**(482), 681–686 (2008)
36. Eurostat: The EU has reached its target for share of persons aged 30 to 34 with tertiary education. https://ec.europa.eu/eurostat/web/products-euro-indicators/-/3-26042019-ap
37. Fellini, I.: Immigrants-labour market outcomes in Italy and Spain: Has the Southern European model disrupted during the crisis? Migr. Stud. **6**(1), 53–78 (2018)

38. Nelder, J.A., Wedderburn, R.W.M.: Generalized linear models. J. R. Stat. Soc. Ser. A (Gen.) **135**(3), 370–384 (1972)
39. Carvalho, C.M., Polson, N.G., Scott, J.G.: The horseshoe estimator for sparse signals. Biometrika **97**(2), 465–480 (2010)
40. Piironen, J., Vehtari, A.: Sparsity information and regularization in the horseshoe and other shrinkage priors (2017)
41. Smith, N.J.: Patsy. https://patsy.readthedocs.io/en/latest/
42. Jackman, S.: Bayesian Analysis for the Social Sciences. Wiley, Chichester (2009)
43. Plummer, M.: Jags: A program for analysis of Bayesian graphical models using Gibbs sampling. In: Proceedings of the 3rd International Workshop on Distributed Statistical Computing, vol. 124, pp. 1–10. Vienna, Austria (2003)
44. Carpenter, B., Gelman, A., Hoffman, M.D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., Riddell, A.: Stan: a probabilistic programming language. J. Stat. Softw. **76**(1), 1 (2017)
45. Stan Development Team: Stan modeling language users guide and reference manual, 2.29. https://mc-stan.org
46. Beraha, M., Falco, D., Guglielmi, A.: JAGS, NIMBLE, Stan: a detailed comparison among Bayesian MCMC software (2021). arXiv preprint arXiv:2107.09357
47. Hoffman, M.D., Gelman, A.: The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. J. Mach. Learn. Res. **15**(1), 1593–1623 (2014)
48. Stan Development Team: RStan: the R Interface to Stan. https://cran.r-project.org/web/packages/rstan/vignettes/rstan.html#see-also
49. Geyer, C.J.: Practical Markov chain Monte Carlo. Stat. Sci. **7**, 473–483 (1992)
50. Giorgetti, I., Picchio, M.: One billion Euro programme for early childcare services in Italy. Metroeconomica **72**(3), 460–492 (2021)
51. Ren, Y., Guglielmi, A., Akhavan, M.: Bayesian inference for investigating gender-based determinants of labour force participation. Submitted

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.