

The pyramid of Okun's coefficient for Italy

Luca Zanin¹

Published online: 27 May 2016
© Springer Science+Business Media New York 2016

Abstract Our aim is to propose a pyramid of Okun's coefficient by age and gender in the Italian labour force using a varying-coefficient model. The unemployment rate by age and gender—useful information for estimating Okun's relationship—is not available for Italy from official statistics. Therefore, we provide an estimation of the indicator using microdata for the 2005–2014 period from ISTAT, the Italian labour force survey. Okun's law is investigated using two measures of the unemployment rate: a traditional measure based on a labour force with and without work experience, and a new measure restricted to the labour force with experience. When Okun's relationship is estimated using the unemployment rate restricted to the labour force with experience, the young population is less sensitive to business cycles. As the workforce ages, this gap in sensitivity tends to shrink. We also found that there are no significant differences by gender in the magnitude of Okun's coefficient among the youngest population when considering the unemployment rate restricted to the labour force with experience.

Keywords Binary generalised extreme value additive model · Economic growth · Okun's coefficient · Unemployment rate · Varying-coefficient model

JEL Classification C14 · C25 · E24 · O40

✉ Luca Zanin
luca.zanin@studio.unibo.it

¹ Prometeia, G. Marconi 43, 40122 Bologna, Italy

1 Introduction

Economic theory suggests that increased production in an economy leads to decreases in unemployment (e.g., Zanin and Marra 2012a). This inverse relationship is well-known as Okun's law. Since the seminal study of Okun (1962), a number of scientific contributions have investigated the relationship between the unemployment rate and economic growth using different methodological modelling approaches and periods of observation (e.g., Viren 2001; Cuaresma 2003; Holmes and Silverstone 2006; Perman and Tavera 2007; Fouquau 2008; Zanin and Marra 2012a). The most common model specification is defined as follows:

$$\Delta UN_t = \alpha + \beta(\Delta GDP_t / GDP_{t-1}) + \epsilon_t, \quad t = 2, \dots, T, \quad (1)$$

where T is the time dimension of the time series, Δ is a difference operator, and UN is the unemployment rate, which is computed as follows:

$$UN = \frac{UN_{NOexp} + UN_{exp}}{UN_{NOexp} + UN_{exp} + Empl} \times 100 \quad (2)$$

where UN_{NOexp} are the unemployed without labour experience, UN_{exp} are the unemployed with labour experience, and $Empl$ are the total employed. The sum of these components measures the labour force. GDP is the country's gross domestic product in real terms, α is the intercept, and ϵ_t is an *i.i.d.* $N(0, \sigma^2)$ random variable. The parameter β represents Okun's coefficient, which is expected to carry a negative sign, based on economic theory.

Recent studies have shown that Okun's coefficient not only is spatially heterogeneous and asymmetric over business cycles but also varies by age cohorts (Hutengs and Stadtmann 2013; Zanin 2014). In particular, these studies suggest that the young population (in particular, the young male population) is the segment that is most exposed to the business cycle. The reasons for this exposure might be attributed to the difficulties faced by the youngest members of the labour force entering the labour market without work experience or to the fact that young people with work experience are frequently employed under temporary contracts and are thereby most exposed to layoffs during recessions or periods of weak economic growth (e.g., Bernal-Verdugo et al. 2013; Agnello et al. 2014; Zanin 2014).

The estimated Okun's coefficient computed by age cohort is a useful measure of the unemployment risk over the life cycle of individuals in relation to the business cycle. It represents an important indicator for economists and policymakers who are interested in establishing effective labour market policies as well as for banks and insurance companies when assessing income risk (for example, in offering solutions regarding insurance coverage, mortgages and loans).

Two main limitations affect the analyses of Okun's law by age cohorts proposed by Hutengs and Stadtmann (2013) and Zanin (2014). First, Okun's coefficient has been estimated using unemployment rates by age cohorts (of 5 years) as derived from national/international statistical office databases. As a consequence, the magnitude of the estimated coefficient is flat within each age cohort. Second, Okun's coefficient is estimated using the unemployment rate (2), which includes

first-time job seekers. Regarding this last point, if Okun's coefficient is used by bank and insurance companies to measure income risk in relation to the business cycle, we expect that this measure will be affected by people who have never received an income from labour and are not eligible for a mortgage or loan. In other words, bank and insurance companies might also be interested in obtaining a measure of Okun's coefficient restricted to those members of the labour force with labour experience. To achieve this aim, we have also estimated Okun's coefficient using the following alternative measure of the unemployment rate:

$$UN = \frac{UN_{exp}}{UN_{exp} + Empl} \times 100 \quad (3)$$

Our purpose is to extend and refine the analysis of Okun's coefficient proposed by Hutengs and Stadtmann (2013) and Zanin (2014). As a case study, we focus on Italy, which is among the most suffering European countries regarding the negative effects associated with the great economic recession related to the financial crisis of 2007–2008. Here, we consider the last 10 years as the time-span for our empirical analysis.

Because measures (2) and (3) by age and gender are not available for Italy from official statistics, we consider a preliminary process to estimate such information using a regression analysis based on microdata from the Italian labour force survey for the 2005–2013 period. Specifically, the idea is to estimate measures (2) and (3) by studying the relationship between the state of unemployment (1 = unemployed; 0 = employed) and the gender and age of individuals. Because the proportion of unemployed individuals in the labour force is typically small (in the 2005–2014 period, the unemployment rate in Italy ranged from a minimum of 6.1 % to a maximum of 12.7 %), the use of a classic probabilistic model that considers a logit link function might not be appropriate. As a possible alternative approach, we consider using an asymmetric link function within a Binary Generalised Extreme Value Additive model (BGEVA) specification, as proposed by Calabrese et al. (2015). In the model specification, we have also relaxed the assumption on the response-covariate relationship by using a penalised smoothing spline approach that allows us to avoid imposing a priori assumptions of linearity or non-linearity on the relationship under investigation. In carrying out a sensitivity analysis, we confirm that the BGEVA model demonstrates better performance compared to the results obtained from models employing a logit or cloglog link function.

After having determined measures (2) and (3) by age and gender, we estimate the relationship between the unemployment rate and economic growth using a varying-coefficient approach within an additive modelling framework, as suggested by Zanin and Marra (2012a, b). This approach allows us to estimate the relationship of interest in a flexible manner. The results obtained from the model's estimation are used to draw the pyramid of Okun's coefficient by age and gender.

The remainder of the article is organised as follows. Sect. 2 describes the data and the econometric methods used in the analysis. The results are discussed in Sect. 3, and the main findings are reported in the final section.

2 Data and econometric methods

In Sect. 2.1, we present the approach used to estimate unemployment rates (2) and (3) by age and gender, and in Sect. 2.2, we introduce the econometric model used to estimate Okun's coefficient by age of the labour force.

2.1 Unemployment rates (2) and (3) by age and gender

To measure unemployment rates (2) and (3) with annual frequency by age and gender, we use microdata from the ISTAT labour force survey for the 2005–2014 period. Here, we consider the labour force between 20 and 64 years of age. In using microdata, we encounter two main constraints: (i) for the surveys from the 2005–2013 period, the age of individuals is provided in classes of age, whereas (ii) for the survey from 2014, the age of the interviewees is known but is not a dimension of stratification. In light of the foregoing, we cannot exclude an issue of robustness in computing unemployment rates (2) and (3) by the age of individuals. To overcome constraints (i) and (ii), we decided to measure unemployment rates (2) and (3) by age and gender using a regression analysis. Specifically, the estimation process was conducted as follows:

1. We defined two sets of sample units for each year of interest: the first set of microdata includes individuals in the condition of UN_{NOexp} , UN_{exp} and $Empl$, whereas the second set includes only individuals in the condition of UN_{exp} and $Empl$.
2. For the 2005–2013 period, for which the age of individuals is provided in classes, we have assumed that the interviewees are uniformly distributed within each class of age.
3. We specified a simple statistical model that allows us to estimate unemployment rates (2) and (3) by age and gender. Specifically, the general model considers the following relationship:

$$UN_i = \alpha + \beta(age_i) \quad (4)$$

where UN_i is a binary variable that takes the value of 1 if an individual is unemployed and 0 otherwise, α is the intercept, age is the age of individuals and β the associated parameter to be estimated. Relationship (4) can be estimated using a classic logit or probit model. Because in our context the proportion of unemployed is typically smaller than the proportion of employed, the estimation of the probability of unemployment using a logit link function may be inappropriate because of its symmetry around 0.5 (e.g., King and Zeng 2001; Calabrese and Osmetti 2013). The estimation of a logit model implies in fact that the outcome curve of the probability of unemployment approaches zero in the same manner as it approaches one. To address this issue, the literature suggests using an asymmetric link function to take into account the imbalance in the proportions of 0 and 1. As mentioned in the introductory section, we use a

BGEVA model specification as proposed by Calabrese et al. (2015). The approach uses a link function based on the quantile function of a GEV random variable. The model is thus specified as follows:

$$\frac{[-\ln(UN_i)]^{-\tau} - 1}{\tau} = \alpha + s(\text{age}_i) = \eta_i \tag{5}$$

where $\tau \in \mathfrak{R}$ is the tail parameter, α is the intercept, and $s(\text{age}_i)$ is a one-dimensional smooth function of the age of the labour force. The smooth function for the covariate age is given as a linear combination of known spline bases, $b_j(\text{age})$, and unknown regression parameters, γ_j . In this way, $s(\text{age}) = \sum_{j=1}^J \gamma_j b_j(\text{age})$, where J represents the number of bases. The number of basis functions determines the maximum possible flexibility allowed for a smooth term. In other words, the higher that J is, the ‘wigglier’ the estimated smooth function will be. Smooth components are typically subject to some identifiability constraints such as $\sum_i s(\text{age}_i) = 0$ (Wood 2006). In Eq. (5), η_i can be written in matrix terms as $\mathbf{B}_i^T \delta$ with $\mathbf{B}_i^T = [1, b_1(\text{age}_i), \dots, b_J(\text{age}_i)]$ and $\delta^T = (\alpha, \gamma_1, \dots, \gamma_J)$.

Replacing the functions $s(\cdot)$ with their regression spline expressions yields a parametric model whose design matrix contains the regression spline bases representing the smooth components in the model. For our case study, we have used the P-splines with 10 basis functions (e.g., Eilers and Marxs 1996; Wood 2006). In principle, Eq. (5) can be estimated using maximum likelihood (ML) estimation. Unfortunately, the classic ML estimation of (5) is likely to result in smooth component estimates that are too ‘wiggly’, hence undermining the utility of the model. This issue can be overcome by using penalised likelihood maximisation.

For a given value of τ , model (5) is estimated by the maximisation of

$$\ell_p(\delta) = \ell(\delta) - \frac{1}{2} \lambda \gamma \mathbf{S} \gamma \text{ w.r.t. } \delta, \tag{6}$$

where $\ell(\delta) = \sum_{i=1}^n -y_i(1 + \tau \eta_i)^{-1/\tau} + (1 - y_i) \ln\{1 - \exp[-(1 + \tau \eta_i)^{-1/\tau}]\}$, $\eta_i = \mathbf{B}_i^T \delta$.

The use of roughness penalties during the model-fitting process is crucial to avoid the problem of overfitting. In expression (6), the smooth has an associated penalty, $\lambda \gamma \mathbf{S} \gamma$, where \mathbf{S} is a positive semi-definite matrix of known coefficients whose values depend on the order of the derivatives (here, set to 2) chosen to represent the roughness of the smooth term, while λ controls the trade-off between the goodness of fit and roughness. The smoothing parameter plays a crucial role. In fact, if λ is too high, then the resulting smooth function will be over smoothed, while if it is too low, then the component will be under smoothed. It is clear that both cases are not ideal, hence the need to use a data-driven smoothing parameter selection procedure (Zanin and Marra 2012a). In our case, parameter estimation is achieved through a generalisation of the

approximate unbiased risk estimator (UBRE, Craven and Wahba 1979; Calabrese et al. 2015). This approach allows for the suppression of that part of the smooth term complexity that has no support from the data (e.g., Wood 2006; Zanin and Marra 2012b; Calabrese et al. 2015). Computations were performed using the `bgeva()` function in `bgeva` package in R (Marra et al. 2015).

4. For a model comparison, we also estimated unemployment rates (2) and (3) by age and gender using a generalised additive model with a logit and cloglog link function (e.g., Wood 2006).
5. We selected the model with the lower value of the Mean Absolute Error (MAE) and Mean Squared Error (MSE). The assessments of model performance were conducted by age cohort and gender of the labour force consistently with the classifications used by national/international statistical offices.

2.2 Estimation of Okun's coefficient by age: a varying-coefficient approach

Okun's coefficient by age and gender of the labour force is estimated by means of the following varying-coefficient model:

$$\Delta UN_{age,t} = \alpha + \beta_{age}(\Delta GDP_t / GDP_{t-1}) + \epsilon_{age,t}, \quad (7)$$

$$age = 20, \dots, 64; \quad t = 2006, \dots, 2014; \quad \beta_{age} = s(age; \gamma),$$

where UN is either annual unemployment rate (2) or (3) as computed in Sect. 2.1, and GDP is the annual real gross domestic product from ISTAT. The vector of real GDP growth effects, $\beta = (\beta_{20}, \dots, \beta_{64})_{age \times 1}$, is modelled as $s(age; \gamma)$. The modifier effect is age . Similarly to model (5) of Sect. (2.1), the use of a smooth term helps to capture the shape of the response-covariate relationship (which is typically unknown *a priori* by the researcher) without imposing any structure on it (e.g., linear or quadratic, using polynomial functions). Specifically, $s(age; \gamma) = \sum_{j=1}^J \gamma_j b_j(age)$, where J represents the number of bases. Here, we have considered 10 basis functions. As described in Zanin and Marra (2012a, b), model (7) is estimated by minimising

$$\|y - X\gamma\|^2 + \lambda \int \{s^d(age; \gamma)\}^2 dage \quad (8)$$

where y is the annual change in the unemployment rate, X is the model matrix containing the basis functions for the varying components interacted with their corresponding GDP growth, γ is the spline parameter vector, the integral measures the roughness of the smooth component, and d is set to 2, indicating the order of the derivative for the smooth term in the fitting process. Considering that regression splines are linear in their model parameters, $\lambda \int \{s^d(age; \gamma)\}^2 dage$ can be re-written as $\lambda \gamma^T S \gamma$, where S is a coefficient penalty matrix. As described in Sect. 2.1, λ controls the trade-off between the goodness of fit and roughness and can be effectively estimated by the minimisation of a prediction error estimate such as the

generalised cross-validation score (e.g., Wo06 and references therein). Computations were performed using the `gam()` function in `mgcv` package in R (Wood 2006, 2016).

3 Results

In this section, we present the results of the estimation process of unemployment rates (2) and (3) by age and gender, and then we present the pyramidal representation of Okun's coefficient obtained after estimating a varying-coefficient model.

3.1 Estimation of unemployment rates (2) and (3) by age and gender

We illustrate the main findings of the estimation process of annual unemployment rates (2) and (3) for the 2005–2014 period by the age and gender of the labour force. From a methodological perspective, we estimated one binary model for each year and gender of the labour force and compared the performance of different link functions: GEV (by assessing the following range of parameters: -0.25 , -0.5 , -0.75 , and -1.0),¹ logit, and cloglog (see Sect. 2). After performing the model estimations, the indicators of MSE and MAE helped us to select the model specification with the best performance (values close to zero). As a result, we found a preference for a GEV link function rather than a logit or cloglog.

Figure 1 shows the estimated unemployment rates (2) and (3). Comparing graphs (a) and (b) with (c) and (d) reveals that the unemployment rate continued to grow during the period under examination, particularly among the youngest members of the labour force. The growth is particularly accentuated when we include in the unemployment rate the people who aspired to enter the labour market for the first time (plots (a) and (b)). This trend is affected by the context of economic uncertainty in recent years, which did not offer the conditions necessary to absorb the labour supply, particularly for the young members of the population who might have just finished their education and were ready to enter the labour market.

Targeting youth unemployment is a priority in the agenda of policymakers because it has several negative implications for society. For example, it can have long-lasting effects on earnings, a negative impact on happiness/well-being and on the long-term choices of individuals (a barrier to obtaining a mortgage/loan or to planning a marriage). For practitioners, the estimated Okun's coefficient can also represent a useful indicator to assess the possible risks and uncertainties associated with the labour market (e.g., loss of income) over the course of an individual's life in relationship to business cycles.

¹ Because we require a curve that approaches 1 rapidly, we assess negative values of τ .

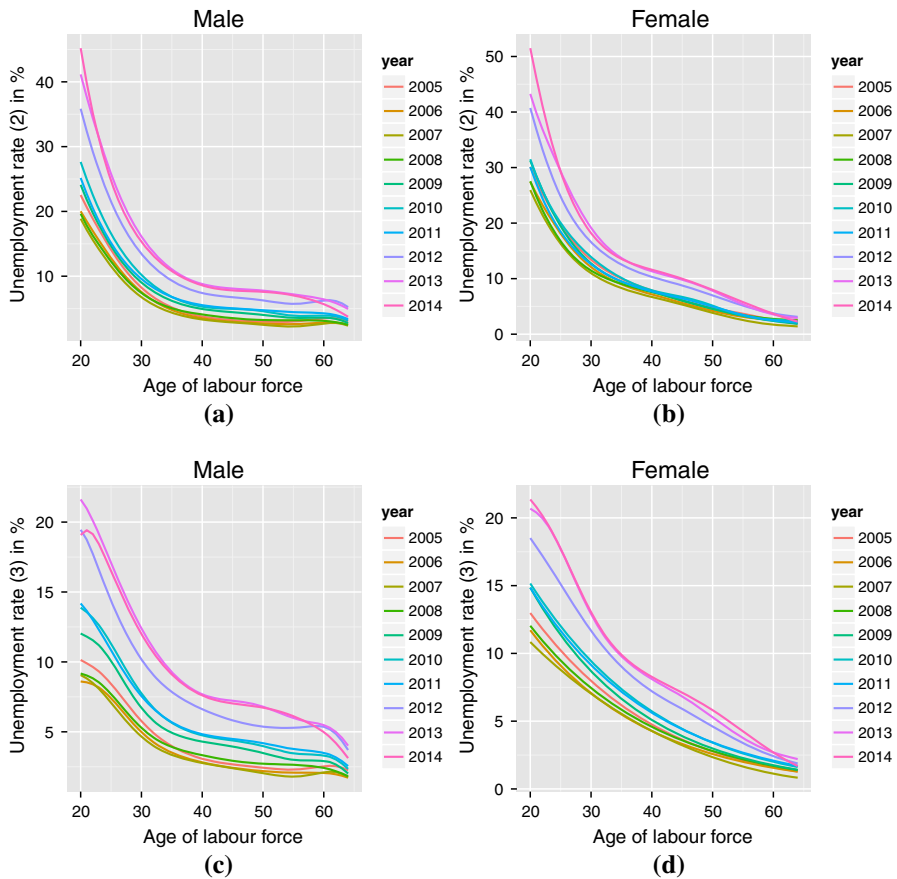


Fig. 1 Plot of the time series of unemployment rates (2) and (3) by age and gender of the labour force

3.2 The pyramid of Okun's coefficient in Italy

Figure 2a, b depict a pyramidal representation of the estimated Okun's coefficients for the 2005–2014 period when unemployment rates (2) and (3), respectively, are employed in equation (7). In comparison with the empirical studies of Hutengs and Stadtmann (2013) and Zanin (2014), the use of a flexible modelling framework allowed us to refine the analysis by the age and gender of the labour force. In other words, we overcome the limitations of the measures of Okun's coefficient that are flat within a fixed age range.

The estimated Okun's coefficient is negative and statistically significant at the 5 % level, with the sole exception of the oldest female labour force (confidence intervals available upon request). As observed by Zanin (2014), we also confirm that Okun's coefficient is variable in magnitude during the life course of individuals. For both male and female members of the labour force, we observe that the estimated Okun's coefficient becomes smaller up to a certain age (approximately 30 years of

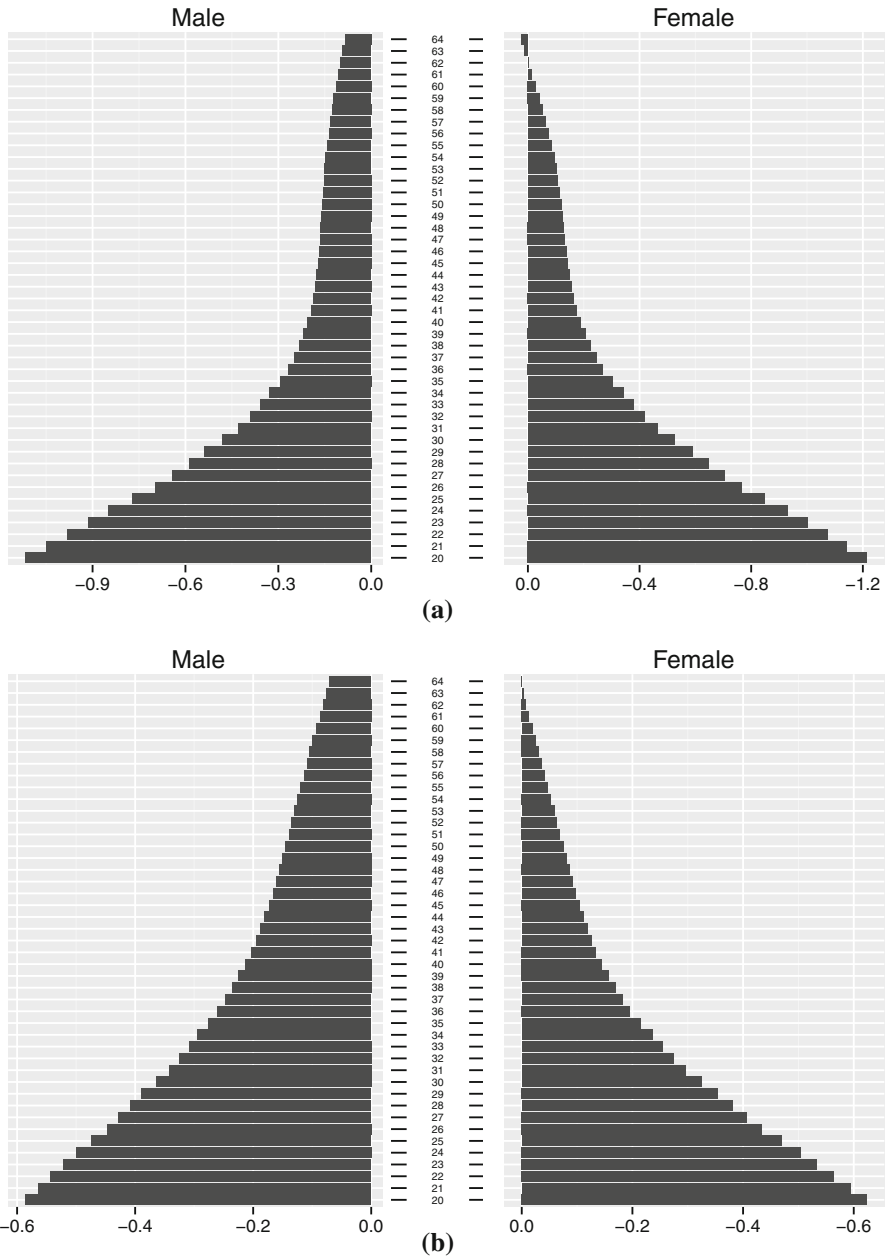


Fig. 2 The Okun's coefficients estimated for males and females for the 2005–2014 period. Plots **a** and **b** report the Okun's coefficients estimated when unemployment rates (2) and (3) are considered, respectively

age) and thereafter tends to stabilise around a certain magnitude. During a negative business cycle, firms and entrepreneurs who aim to increase the number of employees by a few units might impose more stringent selection criteria in searching for the best workers, whereas young workers seek employment that best suits their education. These factors can contribute to frictional unemployment, which may become a serious social problem in the absence of an effective action strategy in policymakers' agendas (see, e.g., Refrigeri and Aleandri 2013; Zanin 2014). Because young people without work experience may encounter greater difficulty in entering labour markets during a period of economic uncertainty/weakness compared with a phase of economic growth, some policies oriented towards vocational training might facilitate the transition from school to work.

Regarding differences between genders, Fig. 2a reveals that the youngest female labour force segment has higher business cycle sensitivity than their young male labour force counterpart. However, as the workforce ages, males becomes more sensitive to the business cycle than females.

Figure 2b represents a novelty in the literature on Okun's law. It shows that when Okun's coefficient is estimated using unemployment rate (3), the youngest segment of the labour force is less sensitive to the business cycle than that observed in Fig. 2a. Moreover, we note a similarity between genders in the sensitivity to business cycles when unemployment rate (3) rather than (2) is employed to estimate Okun's coefficient. For example, for a male and a female of 20 years of age, Okun's coefficient is equal to -1.117 and 1.215 , respectively, when we consider unemployment rate (2), whereas it is equal to -0.586 and -0.624 , respectively, when unemployment rate (3) is considered. In other words, one or more experiences in the labour market for the youngest population can help to significantly reduce the differences between genders in sensitivity to business cycles.

Among the labour force with experience, the higher sensitivity to business cycles of the young labour force than older populations (Fig. 2b) is likely be associated with the use of temporary contracts for the young labour force, thus subjecting the young labour force to a higher probability of layoff (particularly) during periods of economic uncertainty (e.g., Zanin 2014). For a better discussion of this point, the study of a transition matrix (i.e., the probability of being unemployed at time $T+1$, if employed at time T) using a Markov chain framework (see, e.g., Constant and Zimmermann 2014) may be conducted in a future extension of the present study.

4 Concluding remarks

Using data from Italy, we study the relationship between changes in the unemployment rate for male and female age cohorts and economic growth. The data cover the 2005–2014 period on an annual basis. Time series data on unemployment rates by age and gender are not available for Italy from official statistics and must be estimated using a regression analysis. Specifically, as described in Sect. 2, we created two new time series on unemployment rates: one based on a labour force with and without work experience (2), and the other restricted to that segment of the labour force with experience (3).

Using the varying-coefficient model (7) illustrated in Sect. 2.2, we estimate Okun's coefficient by age and gender using unemployment rates (2) and (3). Among the findings of interest, we note that when unemployment rate (3) is used to estimate Okun's law, the youngest population is less sensitive to business cycles than when unemployment rate (2) is used.

The gap in sensitivity between the two estimates [computed as the difference between Okun's coefficient estimated using unemployment rate (2) and Okun's coefficient estimated using unemployment rate (3)] shows a non-linear decreasing trend as the workforce ages. The highest values of the gap in sensitivity are observed among males and females between 20 and 24 years of age (the mean values of the gap (in absolute terms) for males and females are 0.44 and 0.51, respectively).

A further aspect of interest might be to study the pyramid with a focus on the possible asymmetry (over the business cycle) in the magnitude of Okun's coefficient. However, because of data constraints, we were not able to provide evidence in this direction.

The evidence that emerged from our study should encourage policies of transition from school to work to curb the growth of youth unemployment as well as the differences between genders in sensitivity to business cycles. Furthermore, the findings should be of interest for banks and insurance companies in assessing the possible risks associated with the labour market (e.g., loss of income) during the life course of individuals.

Acknowledgments We would like to thank one anonymous reviewer for many suggestions that have helped to improve the presentation and quality of the article. The opinions expressed herein are those of the author and do not reflect those of the institution or affiliation.

References

- Agnello L, Castro V, Jalles JT, Sousa RM (2014) Fiscal adjustments, labour market flexibility and unemployment. *Econ Lett* 124:231–235
- Bernal-Verdugo EL, Fuceri D, Guillaume D (2013) Banking crises, labor reforms, and unemployment. *J Comp Econ* 41:1202–1219
- Constant AF, Zimmermann KF (2014) Self-employment against employment or unemployment: Markov transitions across the business cycle. *Eurasian Bus Rev* 4:51–87
- Calabrese R, Osmetti SA (2013) Modelling small and medium enterprise loan defaults as rare events: the generalized extreme value regression model. *J Appl Stat* 40:1172–1188
- Calabrese R, Marra G, Osmetti SA (2015) Bankruptcy prediction of small and medium enterprises using a flexible binary generalized extreme value model. *J Oper Res Soc.* doi:[10.1057/jors.2015.64](https://doi.org/10.1057/jors.2015.64)
- Cuaresma JC (2003) Okun's law revisited. *Oxf Bull Econ Stat* 65:439–451
- Craven P, Wahba G (1979) Smoothing noisy data with spline functions. *Numerische Math* 31:377–403
- Eilers PHC, Marx BD (1996) Flexible smoothing B-splines and penalties. *Stat Sci* 11:89–121
- Fouquau J (2008) Threshold effects in Okun's law: a panel data analysis. *Econ Bull* 5:9–49
- Holmes M, Silverstone B (2006) Okun's law, asymmetries and jobless recoveries in the United States: a Markov-switching approach. *Econ Lett* 92:293–299
- Hutungs O, Stadtmann G (2013) Age effects in the Okun's law within the Eurozone. *Appl Econ Lett* 20:821–825
- King G, Zeng L (2001) Logistic regression in rare events data. *Polit Anal* 9:137–163
- Marra G, Calabrese R, Osmetti S (2015) Binary Generalized Extreme Value Additive Modelling: package 'bgeva' 0.3. <https://CRAN.R-project.org>

- Okun A (1962) Potential GNP: its measurements and significance. In: proceedings of the business and economics statistics section. American Statistical Association, pp 98–104
- Perman R, Tavera C (2007) Testing for convergence of the Okun's Law coefficient in Europe. *Empirica* 34:45–61
- Refrigeri L, Aleandri G (2013) Educational policies and youth unemployment. *Proced Soc Behav Sci* 93:1263–1268
- Viren M (2001) The Okun curve is non-linear. *Econ Lett* 70:243–257
- Wood SN (2006) Generalized additive models: an introduction with R. Chapman & Hall, London
- Wood SN (2016) Mixed GAM computation vehicle with GCV/AIC/REML smoothness estimation: package 'mgcv' 1.8-11. <https://CRAN.R-project.org>
- Zanin L, Marra G (2012a) Rolling regression versus time-varying coefficient modelling: an empirical investigation of the Okun's law in some Euro area countries. *Bull Econ Res* 64:91–108
- Zanin L, Marra G (2012b) Assessing the functional relationship between CO₂ emissions and economic development using an additive mixed model approach. *Econ Model* 29:1328–1337
- Zanin L (2014) On Okun's law in OECD countries: an analysis by age cohorts. *Econ Lett* 125:243–248