

A Service of

ZBW

Leibniz-Informationszentrum Wirtschaft Leibniz Information Centre for Economics

Bovi, Maurizio

Article The shapley value of age-period-cohort effects

Journal of Applied Economics

Provided in Cooperation with: University of CEMA, Buenos Aires

Suggested Citation: Bovi, Maurizio (2021) : The shapley value of age-period-cohort effects, Journal of Applied Economics, ISSN 1667-6726, Taylor & Francis, Abingdon, Vol. 24, Iss. 1, pp. 297-317, https://doi.org/10.1080/15140326.2021.1932177

This Version is available at: https://hdl.handle.net/10419/314131

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.



WWW.ECONSTOR.EU

https://creativecommons.org/licenses/by/4.0/

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.







Journal of Applied Economics

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/recs20

The shapley value of age-period-cohort effects

Maurizio Bovi

To cite this article: Maurizio Bovi (2021) The shapley value of age-period-cohort effects, Journal of Applied Economics, 24:1, 297-317, DOI: 10.1080/15140326.2021.1932177

To link to this article: https://doi.org/10.1080/15140326.2021.1932177

0

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 04 Aug 2021.

ſ	
L	6

Submit your article to this journal 🖸

Article views: 1906



View related articles



View Crossmark data 🗹



Citing articles: 1 View citing articles 🗹

OPEN ACCESS Check for updates

Routledge

Taylor & Francis Group

The shapley value of age-period-cohort effects

Maurizio Bovi

ABSTRACT

The exact linear dependency among age, period and birth cohort makes it impossible to recover the true parameters of Age-Period-Cohort (APC) models. We then propose to extract reliable information from APC models via the Shapley decomposition, a model-agnostic procedure from game theory that allows to pin down the most likely contribution of each regressor in explaining the variance of the dependent variable. The rationale is that the predicted values of APC models are estimable and the allocation of the R^2 to the APC regressors – interpreted here as the APC "effects" - satisfies desirable properties and produces robust estimates, in that complementing existing methods. We apply the method to the U.S. unemployment rate.

ARTICLE HISTORY

Received 8 June 2020 Accepted 13 May 2021

KEYWORDS

Shapley value; estimability; age-period-cohort models

1. Introduction

Economic behavior and performances often depend on three distinct time effects. People act differently when they are young or old (the age effect). Then, individuals are affected by the current macroeconomic stance (the period effect) as well as by their year of birth (the cohort effect). Age effects pertain to microeconomics in the sense that they reflect the biological and social processes of aging at the individual level, such as employment status, consumption level, marriage, parenthood, retirement, and the like. Period effects emerge at a wider aggregate level, arising from events happening as time passes by that affect all individuals regardless of their age or year of birth. Pandemics, macroeconomic booms and crises, immigration, and important policy reforms are points in case. Birth cohort effects depend on both micro and macro influences. Successive cohorts are differentiated by the changing content of education, peer-group socialization, and idiosyncratic historical experience (Ryder,). For example, Baby Boomers (people born in the two decades after the end of the WWII) could be relatively disadvantaged. They grow up with fewer adults per child, larger class sizes, and fewer entry-level jobs per entry-level job seeker (Easterlin 1978, 1987; Bönke, Corneo, & Lüthen, 2015; Chinhui and McCue, 2017). By the same token, todays' teenagers could suffer from the COVID-19-related educational issues when they will face the labor market in the coming years. For people already in the labor force, instead, the current pandemic is a period effect.

The foregoing suggests that distinguishing Age, Period, and Cohort (APC) effects to address socio-economic processes turns out to be critical. In fact, there is a vast variety and amount of research devoted to APC models. As far as economics, instances can be

CONTACT Maurizio Bovi 🖂 maurizio.bovi@uniroma1.it

*ISTAT - Italian National Institute of Statistics and "Sapienza" University of Rome. The opinions expressed herein are those of the author and do not necessarily reflect the views of the mentioned institutions.

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/ licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited

found in labor economics (Kalwij & Alessie, 2007; Bönke et al., 2015. See also Section 4) as well as in the analysis of savings and consumption (Fukuda, 2006; Schulhofer-Wohl, 2018). But APC studies are also relevant to health economics, international economics, political economy, and industrial organization (McKenzie, 2006; Mehta, Rysman, & Simcoe, 2010; Sang and Seshadri, 2019; Sedláček & Sterk, 2017).

Unfortunately, serious methodological issues affect the statistical analysis of APC models. The crux is the perfect collinearity emerging from the mathematical "age \equiv period-cohort" identity. Since this exact linear dependency makes the model parameters not identified, then they are not the parameters of the model. Actually, they are not the parameters of any model (Greene, 2018). Given the strong interest in disentangling the three APC effects, however, the literature has been proposing various solutions, such as trying to identify APC coefficients using external information (cf. Section 2).

Against this problematic backdrop, our main aim is to suggest a model-agnostic approach to quantify as accurately as possible the APC effects. Since there is no true parameter vector in the unidentified model, we take advantage of the concepts of estimability and Shapley value. Unlike the parameters, in fact, the predicted values of APC models are estimable: no matter the constraint used, the predicted values are always the same (Section 3.1). Hence, statistics based on these predictions - such as the coefficient of determination (R²) - turn out to be equal across APC models. Once computed this common R², we then apply the Shapley decomposition to allocate it among the regressors of the APC model (Section 3.2). This gives the Shapley value (SV) of each regressor. To be sure, the allocation algorithm does not aim to estimate the (unidentifiable) parameters. Rather, it aims to satisfy desirable properties of "fairness" in game theoretical settings that, nonetheless, allows to give the SVs a statistical interpretation (cf. Section 3.2). The computed SVs capture the relative importance or, equivalently, the most likely contribution of each regressor to the APC model's fit, which we interpret as the "effect" of that regressor. It is worth noticing that in APC models the R^2 is typically close to one (Kupper, Janis, Salama, Yoshizawa, & Greenberg, 1983) so that the algorithm distributes virtually all the total variation in the dependent variable and the SVs can therefore be seen as shares of contribution. Also, the SVs are more robust than other ways to measure the relative importance of regressors (such as net effects) because they are computed considering the average of all the models containing that regressor. Finally, Lipovetsky (2012) has shown that they can also be considered elasticities. To put it short, given their importance and due to the impossibility of estimating APC model parameters as well as the difficulty of identify them, we propose to disentangle the APC effects exploiting a procedure that satisfies some desirable properties, does not need external information, and attributes the APC effects with clarity and simplicity.

2. The econometric issue and some existing solutions

Suppose data are organized in a two-way (axp) table with the rows representing categories of age (i = 1,2, ..., a), and the columns defining categories of year of occurrence (j = 1,2, ..., p). Then, define $Y_{ij} = f(R_{ij})$ as a generic function of some rate R_{ij} . In the

empirical part of the paper (Section 4) we use labor market data and $f(R_{ij})$ is the log of the unemployment rate.

The APC model we will focus on has a specific ANOVA-type structure

$$Y_{ij} = \mu + \alpha_i + \beta_j + \gamma_{a-i+j} + \varepsilon_{ij}$$
(2.1)

where μ is the grand mean, α_i , β_j , and γ_{a-i+j} are, respectively, the fixed effect of the *i*th age group, of the *j*th period, and of the cohort associated with the *i*th age group and with the *j*th period. The only random component is ε_{ij} , which is assumed to have distributional properties dependent on the assumptions made about the stochastic nature of Y_{ij} .

It is usual to reparametrize models such as (2.1), expressing each effect as a deviation from the mean of all effects of that type (see, e.g., Yang & Land, 2013):

$$Y_{ij} = \mu^* + \alpha_i^* + \beta_i^* + \gamma_{a-i+j}^* + \varepsilon_{ij}$$
(2.2)

where

$$\mu^* = (\mu + \bar{\alpha} + \bar{\beta} + \bar{\gamma}), \ \alpha_i^* = (\alpha_i - \bar{\alpha}), \ \beta_i^* = (\beta_j - \bar{\beta}), \ \gamma_{a-i+j}^* = (\gamma_{a-i+j} - \bar{\gamma}),$$

and $\bar{\alpha} = \frac{1}{a} \sum_{i=1}^{a} \alpha_i, \ \bar{\beta} = \frac{1}{p} \sum_{i=1}^{p} \beta_j, \ \bar{\gamma} = \frac{1}{a+p-1} \sum_{k=1}^{a+p-1} \gamma_k.$
Hence, $\sum_{i=1}^{a} \alpha_i^* = \sum_{j=1}^{p} \beta_j^* = \sum_{k=1}^{a+p-1} \gamma_k^* = 0.$ This constraint, known as "effect-coding",

implies to have one less effect for age, period, and cohort with respect to the unconstrained case. Centering creates no distortion: results do not depend on which (redundant) age, period, and cohort parameters are removed, and it is better than other reparametrizations (Yang & Land, 2013). Even under this restriction, the APC analysis nonetheless still suffers from the mentioned identification problem.

Equation (2.2) may be written in matrix form as

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\varepsilon}^* + \boldsymbol{\varepsilon} \tag{2.3}$$

where **X** is the design matrix of dimension [ap x 2(a +p)-3]. The left-hand-side vector, **Y**, has dimension [axp], ε^* and ε are, respectively, the vector of coefficients (μ^* , α_i^* , β_i^* , γ_{a-i+i}^*) and of residuals.

Due to the exact relation between age, period and cohort, the matrix (X'X) is not invertible and the least square solution $(X'X)\varepsilon^* = X'Y$ is not unique. The problem is at the population level so that, no matter the amount of data, identification is simply impossible. Though this mathematical confounding cannot be solved by manipulation of the data or the model, it does not mean that there is no way to extract useful information from this kind of data. In fact, the strong interest in the APC effects has stimulated a vast literature. Some of the existing solutions and their drawbacks may be summarized as follows.

One immediate strategy is to perform sub-models, i.e., excluding one of the three APC effects. Often, this is not a good approach. Intuitively, e.g., when considering only age and period effects it is assumed that changes in the dependent variable over time are equal across cohorts, and that period effects play the major role in explaining these changes. In other words, the expected value for the age-period cell is assumed to be only determined by the marginal effects of the i^{th} row (age) and j^{th} column (period), not considering the possible joint effect of the two. Marginalizing likely amounts to impose unnecessary and

possibly unreasonable constraints. Other authors have then imposed less extreme constraints in order to just-identify the model assuming, e.g., that two age, two period, or two cohort parameters are equal. Following Deaton and Paxson (1994), yet another strand of the literature has been imposing that all the linear trends observed in the data can be attributed to the age and cohort effects (Attanasio, 1998; Parker, 1999; Kalwij & Alessie, 2007). That is, the period effects are assumed to be zero mean and orthogonal to a linear time trend. Working on Akaike (1980), then, Nakamura (1986) proposed a Bayesian approach to specify restrictions, whereas the effect parameters are assumed to change gradually. One advantage of this approach is that it does not rule out the possibility of the period effect being orthogonal to a linear time trend (Fukuda, 2006). The main disadvantage affecting all the procedures mentioned so far is that the constraints should be reasonable, which implies one needs to rely on additional information from external sources and/or economic theory. It may turn out to be a hard task (Schulhofer-Wohl, 2018).

The so-called Hierarchical APC (HAPC) approach(Yang and Land, 2008) is based on the underlying idea that time periods and cohort membership represent the social historical context whereas individuals are embedded. This conceptualization is then translated in the model by specifying age as a fixed effect, and period and cohort as random effects. Simulations show that HAPC models are not able to accurately discern APC effects and should be used with caution when there appear to be period or cohort near-linear trends (Bell & Jones, 2018).

As its name suggest, in the proxy variable approach (Heckman & Robb, 1985) one or more proxy variable is used to replace the age, period, or cohort variable in the model. This method shares with those based on constraints the problems of *i*) the search for reliable proxies and of *ii*) proxy-dependent results. Thus, the substitution of age, period, and cohort with measured variables does not necessarily lead to a better model, which opens up to the possibility of having an incorrectly specified model.

Another attempt is the so-called intrinsic estimator (IE, Yang & Land, 2013; Fu, 2018). It is another sort of constrained solution, and it is an application of the Moore-Penrose generalized inverse to the APC problem. It can also be viewed as an extension of principal component analysis whereas the goal is not to reduce data redundancy and to develop predictive models, but to estimate the APC effects. Basically, the logic behind the IE is to remove the influence of the design matrix (which is fixed by the number of age and period groups and not related to the outcome observations) on coefficient estimates. The constraint used by the IE produces estimates that have some desirable statistical properties. For a fixed number of time periods of data, e.g., it has a smaller variance than any constrained generalized linear model estimators. However, the IE is only useful if researchers carefully assess the credibility of the estimates by using theory and side information, and if they keep their conclusions about the effects tentative (Glenn, 2005).

The penalty function setting (Robertson et al., 1999) estimates and compares the results of the three two-effects (AP, AC, PC) and of a (constrained) APC model. The penalty is measured from the differences in the parameters weighted by a measure of goodness of fit and it is used to "identify" the parameters – their values are obtained from the minimization of the penalty. In their study, Robertson et al., (1999) conclude that methods based on the minimization of a penalty function are only of use if the dependent variable is constant over time. The use of goodness-of-fit statistics makes this approach

relatively close to that proposed here. The final procedure that we want to consider here suggests focusing on the identifiable non-linear components of the time effects, leaving the unidentifiable linear components apart (Fannon & Nielsen, 2019). But this latter may contain important information, too. Thought it is the sole similitude, our approach shares with this last procedure the choice to elaborate on estimable objects of APC models (in our case, the coefficient of determination; cf. Chapter 3).

To recap, though imposing constraint(s) or searching for some proxy may be a working identification strategy, the resulting outcomes are completely dependent on which constraint/proxy is chosen. A worth noticing remark shared by all just-identified constrained models, then, is that these latter produce the same levels of goodness-of-fit to the data, which makes it impossible to use model fit as a model-selecting criterion. As we shall show in the next section, finally, solutions based on constraints can be thought of as being unbiased in a peculiar sense – unbiased estimates may be yielded by specifying a constraint that is satisfied by the true parameter values. To take an example, if there are no true cohort effects the age-period model yields a unique set of parameter estimates and may lead to unbiased estimates of the age effects. In such a case, instead, an APC model would give biased and inaccurate estimates. In any case the main, unsolvable, problem is that there is no true parameter vector in an unidentified model such as the APC one.

3. A new approach to apc models: exploiting estimability and shapley decomposition

We now show that *i*) unbiased estimators of the population parameters are only obtained if the chosen constraint perfectly holds among the population parameters, and that *ii*) different constraints lead to different estimates. This paves the way to introduce estimability in our setting. Then, we show that predicted values – hence goodness-of-fit statistics – are estimable in APC models. Finally, we use the Shapley decomposition to estimate the relative importance of each APC effect.

3.1. Estimability in APC Models

Kupper et al. (1983) have shown that the difference between the true (β^*) and the estimated (under some just identifying constraint) parameters ($\hat{\beta}_c^*$) is a function of a vector (V) and the vector of constraints (c', with $c'\hat{\beta}_c^* = 0$) used to compute the estimates:

Bias

$$\widehat{\beta}_{c}^{*} = \mathbb{E}(\widehat{\beta}_{c}^{*}) \beta^{*} = \frac{c' \beta^{*}}{c' V}$$
(3.1.1)

where $\hat{\beta}_c^* = (\hat{\mu}^*, \hat{\alpha}_i^*, \hat{\beta}_j^*, \hat{\gamma}_{a-i+j}^*)$ denotes a set of estimates of β^* , and the vector *V* is the linear combination of the columns of the design matrix (**X**) that produces a zero vector. That is, *V* is the null vector **X***V* = **0**. Since **X** is one-rank deficient, there is only one null vector.

302 👄 M. BOVI

Equation (3.1.1) informs that the bias is zero¹ if $c'\beta^* = 0$. In particular, the bias is zero when **c** is orthogonal to the true parameter vector. Equivalently, one has unbiasedness if the imposed constraint $(c'\hat{\beta}_c^* = 0)$, holds in the population: $c'\beta^* = 0$. Thus, an obvious issue of constrained estimations is that, in order to specify a constraint that is satisfied by the true parameter values, one needs sufficient information about the true parameters. A circular logic may emerge.

Equation (3.1.1) also highlights that different constraints produce different sets of coefficients, underlining the crucial importance of imposing suitable constraints. For instance, Age-Period (AP) models imply over-identifying restrictions (on the cohort effects) that could generate misleading results. A similar logic holds for Age-Cohort (AC) and Period-Cohort (PC) parametrizations, too.

More importantly here, Equation (3.1.1) allows introducing the concept of estimability in our setting.² Parameters are estimable if there is a linear function of the true parameters for which $c'\beta^*=0$. That is, estimable parameters are *i*) invariant with respect to the choice of the constraint employed, and *ii*) unbiased. It can be shown that in APC models predicted values (\hat{Y}) are estimable.

Suppose $\hat{\beta}_{c1}^*$ is a best fitting constrained solution of model (2.3): $\hat{Y} = \mathbf{X}\hat{\beta}_{c1}^*$. Now consider another of the possible constrained best fitting solutions, say $\hat{\beta}_{c2}^* \neq \hat{\beta}_{c1}^*$. Because of the singularity of the design matrix, $\mathbf{X}sV = \mathbf{0}$ (where *s* is a scalar). Hence, $\hat{\beta}_{c2}^* = (\hat{\beta}_{c1}^* + sV)$ is a solution generating the same $\hat{Y}: \mathbf{X}\hat{\beta}_{c2}^* = \mathbf{X}(\hat{\beta}_{c1}^* + sV) = \mathbf{X}\hat{\beta}_{c1}^* = \hat{Y}$.

To sum up, despite there is an infinite number of best fitting solutions, regardless of the constraint used each of these solutions predicts the same \hat{Y} and minimizes the residuals sum of square. Hence, the fit of the model to the observed data is the same for all of the constrained estimators. Finally, notice that the equality of model fit based on different constraints implies that model fit cannot be used to choose between constrained-regression estimators. Estimability, then, shrinks the possibility to find solutions. But it may offer alternative solutions, too.

3.2. Shapley values: computation and interpretation as APC effects

The above digression naturally leads to the next step – to find a procedure to exploit the (estimable) overall goodness-of-fit to extract reliable information from APC models despite the impossibility to estimate the true APC parameters. To this end, we take advantage of the Shapley value (SV). The SV is a solution concept elaborated in cooperative game theory to provide a reasonable allocation scheme of profits obtained by the grand coalition among the players (Shapley, 1956). Thought it is a mathematical framework for sharing the benefit of the cooperation, it is increasingly used in statistics (Lipovetsky & Conklin, 2015; Strumbelj & Kononenko, 2010). Starting from the idea that members of a coalition should receive payments proportional to their marginal contribution, the Shapley value is the fair value (in the sense cleared below) of a player in a cooperative game.

¹The case c' V = 0 is excluded because it implies that the parameters are estimable.

²Cf. Graybill (1976). Parameters are identifiable if there is an estimator that would produce the true, underlying datagenerating parameters on a sample of infinite size. Thus, identification implies estimability. Estimability, instead, does not imply identification (Christensen, 2011).

JOURNAL OF APPLIED ECONOMICS 😔 303

Formally, a cooperative game consists of a set of N players and a real-valued (characteristic) function, v, that maps a value v(S) to each coalition $S \subseteq N$, with $v(\emptyset) = 0$. Assuming that all players cooperate, the question is how to split the value of the grand coalition v(N) among the participants. The Shapley value (SV) offers an answer. The SV of player $i \in N$ is defined as

$$\phi_i(\nu) = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(|N| - |S| - 1)!}{|N|!} [(\nu(S \cup \{i\}) - (S)]$$
(3.2.1)

The SV can be interpreted as a weighted average of the marginal contribution, $[(v(S \cup \{i\}) - v(S)], of element \{i\} in all combinations. Specifically, the SV allocated to player$ *i* $is based entirely on this marginal value that this player contributes when joining each coalition <math>S \subseteq N \setminus \{i\}$. As per the weight, suppose the players form the coalitions by joining, one-at-a-time, in the order defined by a permutation π of N. That is, player *i* joins immediately after the coalition $S_{p,i} = \{j \in N : p(j) < p(i)\}$ has formed, contributing marginal value $[v (S_{p,i} \cup \{i\}) v(S_{p,i})]$. Then, the SV is the average marginal value contributed by player *i* over all |N|! permutations π , that is,

$$\phi_{i}(\nu) = \frac{1}{|N|!} S_{p}[(S_{p,i} \cup \{i\}) - (S_{p,i})]$$
(3.2.2)

The equivalence of Equations (3.2.1) and (3.2.2) is due to the fact that |S|!(|N| - |S| - 1)! is precisely the number of permutations π for which $S = S_{p,i}$, since there are |S|! ways to permute the preceding players and (|N| - |S| - 1)! ways to permute the succeeding players.

A key advantage of the Shapley value is that it is the only attribution method that satisfies all (i.e., all other methods fail to satisfy at least one) of the following nice properties:

- (1) *Efficiency* The grand coalition value should be entirely allocated among the players:
- $\nu(N) = S_{i \in N} f_i(\nu)$
- (1) *Symmetry* If two players are substitutes in the sense that they contribute the same to each coalition, the solution should treat them equally:

if $v(S \cup \{i\}) = v(S \cup \{j\})$ for all $S \subset N \setminus \{i, j\}$ then $\phi_i(v) = \phi_j(v)$

(1) Linearity/Additivity – If v, v' are two games with the same set of players N, then

 $\phi_i(\alpha v + \alpha' v') = \alpha \phi_i(v) + \alpha' \phi_i(v')$ for all $\alpha, \alpha' \in \mathbb{R}$

(1) *Dummy player* – If a player contributes nothing to any coalition, then the allocation should attribute him nothing:

if $v(S \cup \{i\}) - v(S) = 0$ for all $S \subseteq N \setminus \{i\}$, then $\phi_i(v) = 0$

To insert the SV in the APC setting, suppose to estimate by OLS the APC model of equation 2.3 imposing that, e.g., the first two age effects are equal. As shown in Section 3.1,

the obtained $R^2(R_N^2)$ would not change using other restrictions.³ R_N^2 is then seen as the grand coalition value of a cooperative game (the APC model), where the players (the right-hand-side variables) work in coalition to explain the phenomenon (the dependent variable).

The SVs result from a particular allocation of the model's fit among the regressors, one which considers all possible coalitions/regressions, and it is fair in the sense that it meets the above cited properties. To be sure, they are not statistical properties. But they nonetheless allow to pin down the most likely contribution, or relative importance, of each regressor in the model (more on that below). This contribution is what we interpret as its "effect". The statistical reading of the properties of the algorithm in the present context is as follows. The R_N^2 should be entirely allocated among the regressors (Efficiency); if two regressors contribute the same to each regression, then it sounds natural to require that their relative importance be the same (Symmetry); if the gains of two coalitions/regressions are combined, then the distributed gains should correspond to the sum of the two individual gains (Linearity/Additivity); if an independent variable contributes nothing to any regression, then it is desirable that its explicative power be zero (Dummy player). Note, finally, that the SV is the average contribution of a regressor to the coefficient of determination considering all regressions, not the difference in the R² when we remove a regressor from the model.

The following example on the practical computation of the SVs in the present framework helps to make progress. Consider N = 4 regressors⁴ x_i (i = 1, ..., 4.). Let $R^2_{**\setminus x_i}$ be the R^2 relative to all the possible models with two regressors excluding x_i , and $\bar{R}^2_{**\setminus x_i}$ be its mean value. Similarly, define $\bar{R}^2_{x_i**}$ as the mean value of the R^2 relative to all the possible models with three regressors one of which is x_i . The SV for regressor x_i may then be computed as:

$$\boldsymbol{\phi}_{x_i}(\boldsymbol{\nu}) = \frac{1}{N} \left[(R_N^2 - \bar{R}_{***\backslash x_i}^2) + (\bar{R}_{x_i**}^2 - \bar{R}_{**\backslash x_i}^2) + (\bar{R}_{x_i*}^2 - \bar{R}_{*\backslash x_i}^2) + R_{x_i}^2 \right]$$
(3.2.3)

where R_N^2 is the goodness of fit of the "all-in" regression, and $R_{x_i}^2$ refers to the regression with only x_i as right-hand-side variable (apart from the always present constant). Therefore, the SV for a predictor x_i is a measure of its importance in explaining the APC model's fit or, in other terms, the Shapley decomposition imputes the most likely contribution of each individual regressor to the overall R_N^2 of an APC model. The crux is that some of the regressors may have large effect on the APC model's fit, while others may be irrelevant. Thus, the SV considers what the fit would have been if that regressor was absent; the bigger the change in the R_N^2 , the more important is the regressor. Observing only a single regressor at a time, however, disregards the dependencies between it and the other regressors, which produces inaccurate and misleading outcomes in APC analyses (and not only). But notably, the SV takes into account how the fit changes for each possible subset of independent variables. The Shapley procedure finds the marginal contribution of each regressor in all the possible combinations/regressions; thus, it is a reliable way to disentangle the contributions to the R_N^2 of each right-hand-side variable. In so doing, then, it is similar to the regression anatomy of a coefficient, i.e., the bivariate slope coefficient after partialing out all other regressors in a multivariate model. Comparing the SV to the net effect points out other statistical properties of the SV.

³Table 1 shows that using other restrictions do not affect the APC model's fit (cf. Chapter 4).

⁴Although not explicitly indicated, the intercept is present in all the models.

The net effect of a regressor depends on its direct (as measured by its coefficient squared) and indirect (measured by the combination of its correlations with other variables) effects. Though net effects satisfy the efficiency property (the sum of the net effects equals the model's coefficient of multiple determination), they can be subjected to the multi-collinearity in the data so that the estimated net effects can be negative, which is difficult to interpret. The proposed SV, instead, is always positive: any additional variable increases the quality of data fitting. More importantly, relative importance measures based on the SV are more robust than net effects because the SV is an average across all possible models with different subsets of the regressors. Finally, Lipovetsky (2012) has showed that SVs can be interpreted as elasticities as well.

So far so good. A possible disadvantage of the Shapley procedure is in the exponentially increasing number of regressions to estimate: the number of all possible models is equal to 2 ^N-1. As we shall see in the next section, with six groups of age and six periods, which gives eleven cohorts, we have N =23 for a total of 8,388,607 regressions. This said, however, computing speed is continuously increasing.⁵

In summary, taking advantage of the SV allows to discern the most likely contribution – the relative importance – of each individual regressor to the R^2 of an APC model. We refer to the magnitude of the influence of each right-hand-side variable as the corresponding APC effect. Since the R^2 is typically very close to one in APC models (cf. Introduction), in the present setting the SVs may also be interpreted as shares of contribution.

4. Empirical application. The U.S. unemployment rate

In this section, we apply the Shapley decomposition to quantify the age-period-cohort effects in an APC model of the U.S. unemployment rate. In so doing, we also offer some empirical evidence of how the constraints affect the results in APC analyses (cf. Table 2).

Data for unemployment rates are from the OECD's online database, cover the period 1960–2017, and are computed as the number of unemployed people as a percentage of the labor force. In this application, thus, the APC effects represent three distinct ways in which the employment status of an individual can change over time. For instance, the age effect emerges if there are differences in the unemployment rate of different demographic groups regardless of which period and cohort is considered. The estimation problem is that we cannot observe the employment status of a single person with different ages at a single point in time to isolate the age effect. Similarly, we cannot identify period effects – it is not possible to observe a single person with the same age at two different points in time. In a cross-section study an unemployment trend decreasing with age may be due to cohort effects that are not accounted for. The old in a cross section may not be more employed than the middle aged because they are old but, e.g., merely because they belong to less numerous cohorts. The availability of panel data does not really solve the problem because of confounded age and time effects (which was not the case in a cross section). Following one cohort over time one can never be fully sure whether the change in the employment status is due to the aging process or to passage of time. It is therefore problematic to discuss age effects without

⁵After programming efforts, nonetheless, even ordinary PCs allow this kind of computations in an acceptable time span. We have used the software RStudio, and a basic PC equipped with a dual-core CPU, 12gb RAM and 3.1 GHz.

considering cohort and time effects at the same time. Moreover, all three of them might be important for both the positive and normative viewpoint so it is not warranted to dismiss either one of the three a priori. To take an example, the importance of the generational aspects of unemployment relative to life cycle (age) and business cycle (period) impacts suggests that policies should address the structural issues affecting each of the birth cohorts, rather than focusing on age groups per se. Likewise, a significant period effect would suggest that interventions would be worthwhile for individuals at all stages of the life course.

4.1. Data and Preliminary Analysis

Since data for age is recorded in ten-year intervals, for illustrative purposes we aggregate the data to get an Age-Period dataset with base unit ten. The aggregation leaves us with 6 periods, 6 ages and 11 cohorts. Due to the aggregation of the age groups, cohorts can only be imprecisely computed. We take the mid-age of each group to compute the eleven 10-year cohorts which start from 1890 to 1899.

The last line of Table 1 informs on the dramatic rise between 1960s and 1980s in U.S. total unemployment. Since then, things changed. The 1990s opened with a brief recession that was followed by a sustained decline in unemployment that then stabilized at lower levels until the late 2000s. In the latest decade under scrutiny the Great Recession of 2007–2009 and the subsequent slow recovery pushed, with the usual lag, the unemployment rate to levels as high as 6.8 per cent. The point here is that examining aggregate data may hide several potentially critical details.

Inspecting unemployment by age one may wonder about the role of aging for the level of unemployment. Age groups differ in their employment-related attributes – such as productivity, matching efficiency, and labor turnover – that are independent of cohort and period. From Table 1 one may easily reckon that the unemployment rate of the youngest is typically twice than the overall rate, with only minor oscillations. This kind of information is more easily visible in Figure 1 that displays the age-specific unemployment rates by time period.

Each period-line depicted in Figure 1 offers a cross-sectional look at the rates by age. It stresses the presence of decreasing trends and a noticeable age effect: no matter the period under scrutiny, once in the labor market young people face significant challenges in finding job. Figure 1 clearly shows that youngsters have been suffering from structurally higher unemployment with respect to all the other demographic groups – since the 1960 the unemployment rate of individuals aged 15–24 is always well above that of the others.

Period						
Age	1960-69	1970-79	1980-89	1990-99	2000-09	2010-17
15-24	9.5%	12.7%	13.9%	12.0%	11.9%	14.0%
25–34	4.0%	5.6%	7.1%	5.7%	5.6%	7.1%
35–44	3.3%	3.9%	5.1%	4.4%	4.4%	5.5%
45–54	3.1%	3.5%	4.4%	3.6%	3.9%	5.2%
55-64	3.3%	3.3%	4.1%	3.6%	3.8%	4.9%
65–74	3.7%	4.2%	3.2%	3.5%	3.9%	5.1%
Total	4,6%	6,3%	7,2%	5,7%	5,6%	6,8%

Table 1. U.S. Unemployment Rate by Age.

Source: OECD's online database.



Figure 1. U.S. Unemployment Rate: Age Trend by Period. Note: Unemployment rates (vertical axis) by age group (horizontal axis).

Many authors have elaborated on that. Shimer (Shimer, 1998) has claimed that aging of the labor force is important in explaining the decline in unemployment. Others have argued that youth unemployment is rooted in long-standing structural obstacles that prevent many youngsters from making a successful transition from school to work (International Labour Organization (ILO), 2017; Kelly, McGuinness, & O'Connell, 2012). Many talented young people are still pursuing their education at these ages, so the level of formal education of the youth in the labor market is lower than in older age groups. Productivity may then increase with age when job experience is important (Autor, Levy, & Murnane, 2003). Finally, the economic literature has been stressing that the youth labor market has undergone perverse structural changes in recent decades. Labor demand has shifted away from routine work and towards jobs that require technical skills or post-secondary training (Acemoglu & Autor, 2011). As per the supplyside, alternative sources of labor such as adult middle-skill workers or immigrants may be filling jobs traditionally held by youth (Smith, 2011, 2012).



Figure 2. U.S. Unemployment Rate: Period Trend by Age. Note: Unemployment rates (vertical axis) by age group throughout six decades (horizontal axis).

Figure 2 helps to consider the period effect, i.e., similar changes in the unemployment of all individuals at a point in time due, e.g., to peculiar macroeconomic stances. With the exception of the oldest group, Figure. 2 highlights a common trait shared by all the demographic groups - the relative peaks recorded in the 1980s and in the latest years of the sample. The deep crises of these periods (triggered by, respectively, the 1979 oil shock and the 2007–2009 financial crisis) seem to have had a widespread – and, as usual, lagged – impact on unemployment. This almost regardless of age. Comparing the various demographic groups, however, unemployment cyclicality appears to shrink with age. Figure 2 points out that, to a larger extent than the older groups, the employment of under 35 years old Americans appears to act as a cyclical buffer, rising substantially during booms and falling dramatically amid crises. The fluctuations recorded by under 25 people are even more accentuated. Among the reasons may be that young have more flexible working arrangements and that, during bad times, they are the first to be fired given their lack of work experience and firm-specific human capital. (Bell and Blanchflower 2011; Dietrich and Möller 2016). It may also be that crises push individuals into the labor market prematurely and without the adequate skills to be structurally absorbed into the workplace.

The previous figures do not explicitly account for possible cohort effects. To this end, Figure 3. displays unemployment rates against cohorts. In this plot each curve represents one age group, in that offering yet another useful viewpoint to examine the U.S. unemployment rate. On the one side, the picture is consistent with the presence of some cohort effects, i.e., differences in unemployment rates only due to the year of birth. Data show that same-age individuals born in more recent years suffer from more unemployment with respect to older generations. When they were 15–24 years old, e.g., people born in the 1960s or in the 1990s faced unemployment rates as high as 14%, much higher than those of same-age agents born in the 1940s. At least in part, however, these differences might also be due to period effects:



Figure 3. U.S. Unemployment Rate: Cohort Trend by Age. Note: Unemployment rates (vertical axis) by age groups and eleven (year-of-birth) cohorts (horizontal axis).

when they were 15–24, Americans born in the 1940s faced less dramatic labor market stances than other demographic groups. On the other side, Figure 3 confirms the importance of age in the U.S. labor market, with the line representing people aged 15–24 having much higher levels than the others. No matter your birth cohort if you are young, you have less chances to work with respect to older people. Age seems therefore to impinge on youth unemployment more than factors such as the decreasing fertility rate that, reducing the size of today youth cohorts lessens the competition in the labor market.

All in all, U.S. unemployment rates seems to vary over the three time dimensions albeit with different sizes. Specifically, the phenomenon appears to show a dominance of age effects with respect to cohort and period effects. But there are logical reasons sustaining all the three effects and evidence based on a graphical display of data cannot be satisfying. A more refined analysis is thus needed to better qualify and – even more so – quantify the relative contribution of the time trends under analysis.

4.2. More Formal Analyses and APC Effects as Shapley Values

Table 2 considers six alternative models for the log of the unemployment rate: Age (only age effects), Period (only period effects), Cohort (only cohort effects), AP, AC, and "all

Model	Age	Period	Cohort	Age-Period	Age-Cohort	APC (Constraints)		
				5	5	IE	$a_1^* = a_2^*$	$\beta_2^* = \beta_3^*$
Intercept	-2.97*	-2.97*	-2.92*	-2.97*	-2.97*	-2.97*	-2.97*	-2.97*
15–24	0.87*	_	-	0.87*	0.71*	0.78*	-1.05*	0.52*
25–34	0.12*	-	-	0.12*	0.02	0.05*	-1.05	-0.11
35–44	-0.15*	_	_	-0.15*	-0.19*	-0.18*	-0.55*	-0.23*
45–54	-0.27*	_	_	-0.27*	-0.23*	-0.23*	0.14*	-0.18*
55-64	-0.30*	_	_	-0.30*	-0.20*	-0.23*	0.87*	-0.07
65–74	-0.27	-	-	-0.27	-0.12	-0.19	1.64	0.07
1960–69	-	-0.22	-	-0.22	-	-0.14*	1.69*	0.12
1970–79	-	-0.05	-	-0.05	-	0.02	1.12*	0.17*
1980–89	-	0.08	-	0.08*	-	0.12*	0.49*	0.17*
1990–99	-	-0.04	-	-0.04	-	-0.06*	-0.43*	-0.12*
2000-09	-	-0.01	-	-0.01	-	-0.08*	-1.17*	-0.23*
2010–17	-	0.23	-	0.23	-	0.14	-1.69	-0.12
Coh 1890–99	-	_	-0.38	-	-0.20	0.01	-3.65*	-0.51*
Coh 1900–09	-	_	-0.37	-	-0.17	-0.06	-2.98*	-0.47*
Coh 1910–19	-	_	-0.52*	-	-0.29*	-0.25*	-2.45*	-0.56*
Coh 1920–29	-	_	-0.41*	-	-0.18*	-0.14*	-1.60*	-0.34*
Coh 1930–39	-	_	-0.31*	-	-0.12*	-0.08*	-0.81*	-0.18*
Coh 1940–49	-	_	-0.05	-	0.01	2.E-5	2.E-5	2.E-5
Coh 1950–59	-	_	0.10	-	0.12*	0.08*	0.81*	0.18*
Coh 1960–69	-	_	0.19	-	0.16*	0.10*	1.56*	0.31*
Coh 1970–79	-	_	0.28	-	0.16*	0.12*	2.31*	0.43*
Coh 1980–89	-	_	0.53*	-	0.21*	0.13*	3.06*	0.54*
Coh 1990–99	-	-	0.95	-	0.29	0.09	3.74	0.60
Adj.R ²	0.85	-0.06	0.42	0.95	0.91	0.98	0.98	0.98
R ²	0.87	0.09	0.59	0.97	0.95	0.99	0.99	0.99

Table 2. Estimable and non-estimable statistical objects in APC models.

The APC model is $Y_{ij} = \mu^* + a_i^* + \beta_i^* + \gamma_{a-i+j}^* + \varepsilon_{ij}$ (cf. equation 2.2), where dependent variable = Log(unemployment rate). All regressions are performed constraining the sum of the coefficients of age (a_i^*) , of period (β_i^*) and of cohort (γ_{a-i+j}^*) to be zero (that is, we have used "centralization"). As per the APC models, the additional restrictions are: $(a_1^* = a_2^*)$, i.e., the coefficients of the first two age effects are equal; $(\beta_2^* = \beta_2^*)$, i.e., the coefficients of the 2nd and 3rd period effects are equal; IE = Intrinsic Estimator constraint. "*" means p-value ≤ 0.1 . Bold coefficients indicate the reference factor. in" APC models.⁶ The main aim here is to stress the change in the estimated parameter set when using disparate constraints and, possibly more important for our aim, the estimability of goodness of fit statistics.

Table 2. confirms that the estimated parameters depend on the (unavoidable) constraint imposed. It also underlines the peculiar structure of the APC data. While the AP model has the same coefficients of its two sub-models, namely the models A and P, the coefficients of the AC model are different from those of the A and C models. This is due to the fact that in Age-Period tables such as Table 1, only cohorts generate interactions effects – restricting to zero the cohort effects mechanically "separates" the other two effects. But these zero restrictions may be misleading. Similar considerations apply when estimating an AC model, which amounts to impose no period effects.

A paramount indication of Table 2 is that, though several models fit the data quite well, the highest adjusted R^2 are those of the APC models. In fact, it suggests that one should continue with this all-in parametrization (Yang & Land, 2013). As expected, then, reflecting the estimability of the predicted values all the estimated APC models have the same R^2 regardless of the constraint used. Another as expected as critical outcome is that these all-in parametrizations suffer from the identification problem and, as shown by the results collected in the last columns of Table 2, the estimated parameter vector changes according to the constraints used to compute it. How, then, to quantify the APC effects?

Taking stock of the above reported information and the evidence gathered in the previous section it turns out that we are in the case whereby the use of the proposed Shapley allocation procedure is advisable. Figures 3, 4 and 5 summarize the APC effects as quantified by the Shapley values.

Before looking at the results in detail is perhaps worth recalling that we have twentythree SVs, six for age and period plus eleven for cohorts. The Shapley algorithm computes these values distributing among them the coefficient of determination of the APC regression. The allocation has the statistical interpretations and the desirable properties discussed in Section 3. According to Property 1, e.g., the sum of the twentythree SVs reported in the above mentioned three figures necessarily equals the value of the R² to be distributed which, as reported in Table 2, is 0.99.

Figure 4 confirms and quantifies the extreme difficulty for under 25 years old individuals in the U.S. labor market. The corresponding age effect accounts for more than 40 per cent of the R², which is an impressive number considering that the procedure must "fairly" distribute 0.99 to twenty-three players/effects. This large amount is nonetheless unsurprising in view of the literature discussed as well as the preliminary evidence reported in the previous section. The age effect does not decrease smoothly, however. Although their size is just above 5 percent – i.e., about eight time smaller than that of the youngest – in fact, the second and third largest most likely contributions refer to individuals aged, respectively, 45–54 and 55–64. Labor economists have offered sound explanations that can be summarized by saying that this kind of unemployed is characterized by significant difficulties in finding a job. Some authors, e.g., have pointed out that people at this stage of the life cycle tend to have a reservation wage that is too high

⁶We do not estimate the Period-Cohort model because, as suggested by the preliminary evidence gathered in Section 4.1, age effects are definitively important in the APC analysis of US unemployment.



 $\phi_{x_i}(v) = \frac{1}{N}[(R_N^2 - \overline{R}_{x**/x_i}^2) + (\overline{R}_{X_{i**}}^2 - \overline{R}_{x*/x_i}^2) + (\overline{R}_{X_{i*}}^2 - \overline{R}_{x*/x_i}^2) + (\overline{R}_{x*/x_i}^2 - \overline{R}_{x*/x_i}^$ value. Similarly, define $R_{X_{y**}}^2$ as the mean value of the R^2 relative to the models with three regressors one of which is x_i . The SV for regressor x_i is computed as: Figure 4. U.S. Unemployment: The Shapley Value of the Age Effect. The numbers around the line represent the most likely contribution of each age to the overall R2 of the APC model. This contribution is the age "effect", and it is the Shapley value (SV) that is computed as follows. Apart from the always present constant, consider for simplicity N = 4 regressors x_i (i = 1, . . , 4). Let R_{**/x_i}^2 be the R^2 relative to all the possible models with two regressors excluding x_i , and R_{**/x_i}^2 be its mean the regression with only x_i as right-hand-side variable.



Figure 5. U.S. Unemployment: The Shapley Value of the Period Effect. The numbers around the line represent the most likely contribution of each period to the overall R2 of the APC model. This contribution is the period "effect", and it is the Shapley value that is computed as described under Figure 1.





(Axelrad, Luski, & Malul, 2017); others that these agents find it difficult to keep up in fastchanging technological environments (International Labour Organization (ILO), 2017).

Figure 5 points out that the most likely contribution of the calendar time – i.e., the period effect – is always rather small. There are only two episodes where the percentage of the coefficient of determination allocated by the Shapley procedure to the period effect is (marginally) above the 2 per cent, one at the beginning, the other at the end of the sample. This is an expected result in light of what said in Section 4.1. But it may also be due, at least in part, to the fact that in our data time is aggregated in decades (as data for age is recorded in ten-year intervals) which mechanically dilutes the effects of the cyclical factors that are among what the period effects try to capture. It is then not astonishing that the SVs apportioned by the proposed procedure to the calendar year result so small or, equivalently, that period effects explain such a trivial portion of the U.S. unemployment rate.

Finally, the SVs collected in Figure 6 sustain the previously cited visual impression that individuals born in the 1940s show the smallest cohort effect. In fact, the algorithm ends up assigning to it as little as 0.5 percent. Unlike what happen at other cohorts under scrutiny, in other terms, for these Americans the year of birth is not an unfavorable starting point for their employment status. A somewhat intriguing outcome emerges comparing these individuals, born during the WWII, to those born during the previous world war (cohort 1910-1919). Actually, Figure 6 informs that these "WWI-born" Americans contribute to the explanation of the U.S. unemployment rate much more than "WWII-born" Americans. We leave indications about the causes in the agenda but for our main aim we observe that in the sample period under scrutiny (1960-2017) these unemployed are never under 35 years old people. It seems thus correct that the proposed technique attributes part of their unemployment status to a relatively high cohort effect. As per the generations born after the WWII, finally, the Shapley decomposition emphasizes a dramatic ever-worsening situation. This apparently unstoppable rising trend has led the youngest generation to contribute 6 percent to the explanation of the APC model's fit, which marks an historic record as far as cohort effects. Among the possible explanations it may be cited the massive expansion of higher education at a faster pace than the demand for educated workers that has increasingly raised the competition in the labor market for the younger, better-educated, cohorts. What is surer, as well as important, is that the escalating cohort effect detected by the SVs is a clear normative signal for U.S. policymakers on the growing difficulties faced by the youngest generations of American citizens in the labor market.

5. Conclusions

Age, period, and year of birth are linked by a mathematical linear relationship that makes it impossible to recover the true parameters in APC models. The substantive interest in the information potentially contained in these models has led the literature to suggest several solutions. Among them, the identification problem has often been addressed by resorting to constraints and external information. But this is not always possible, and in any case, results are as reliable as the constraints and the outside information are.

We have then suggested a novel approach that, keeping away from the mission impossible to recover the unidentifiable parameters of APC models, aims nonetheless to extract as reliable as possible information from the models without using constraints and external information. The proposal consists in applying a model-agnostic procedure from game theory, the Shapley decomposition, to the goodness-of-fit statistic of the APC model. Unlike the parameters, in fact, these statistics are estimable in the APC setting. Specifically, we have used the R^2 of the APC model as the grand coalition payoff of a coordination game played by the APC regressors. Then, we have used the Shapley decomposition to allocate the R^2 to the regressors according to the relevance of these latter. The resulting Shapley values (SVs) are here interpreted as the APC effects.

Besides pinning down the most likely contribution of each regressor in explaining the variance of the dependent variable, the procedure has several advantages. Since the coefficient of determination of APC models is typically very close to one, the computed SVs are shares of relative importance of the regressors. The Shapley algorithm, then, is unique in satisfying several desirable properties such as efficiency, symmetry, additivity, and dummy player. In addition, the Shapley allocation method is more robust than other ways to disentangle the relative importance of regressors (such as net effects), and attributes the APC effects with clarity and simplicity. All in all, it complements existing methods to approach the APC models.

We have applied the proposed approach to estimate an APC model of the U.S. unemployment rate. Our main findings can be summarized as follows. As expected, the regressors explain virtually all the variance of the unemployment rate ($R^2 = 0.99$). Emerge, then, a dominant age effect. Specifically, the procedure allocates 41 percent of the coefficient of determination to the first age group (15-24) regressor. This is a really remarkable effect considering that almost halves the percentage of R² that must be fairly divided among the remaining twenty-two regressors. Remarkable but not unexpected. The age effect is consistent with the wide strand of the literature stressing the longstanding structural obstacles faced by young Americans in the labor market. Though to a much smaller extent, other two problematic demographic groups in U.S. labor market turn out to be individuals aged 45-54 and 55-64 which, together, obtain from the procedure more than the 10 percent of the R². It suggests that the age effects that applies across all cohorts and periods are not linear. Period effects are almost absent perhaps because data are aggregated in decades which, of course, hampers the importance of the macroeconomic fluctuations that these effects should capture. This notwithstanding, there is information in these outcomes, albeit about the procedure rather than U.S. unemployment. Actually, it seems to be fair to distribute small quantities of explicative power to a player that mechanically cannot participate that much to the explanation. Instead, there is some larger cohort effects, i.e., inter cohort differences in the unemployment rate net of age and period effects. Contrasting individuals born during the two world wars evidence shows that the cohort effect for Americans born in the decade 1910-1919 is much larger than that of Americans born in the decade 1940-1949. The very small SV allocated to this latter cohort (0.5 percent), in fact, marks the lowest point of an apparently bulletproof upward trend that has made the most recent cohort here studied (1990-1999) the most disadvantaged one, displaying a contribution of 6 percent. This ascending trend is congruent with the increasing competition in the labor market for the younger due to the massive expansion of higher education at a faster pace than the demand for educated workers.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

- Akaike H. (1980). Likelihood and the Bayes procedure. In Bayesian Statistics, Bernardo JM, DeGroot MH, Lindley DV, Smith AFM (eds). University Press: Valencia.
- Acemoglu, D., & Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. In O. Ashenfelter & D. E. Card (Eds.), *Handbook of Labor Economics* (Vol. 4(B), pp. 1043–1171), Elsevier, Netherlands.
- Attanasio, O. (1998). Cohort Analysis of Saving Behavior by U.S. Households. *The Journal of Human Resources* 33(3,: 575-609.
- Autor, D., Levy, F., & Murnane, R. (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly Journal of Economics*, 188(4), 1279–1333.
- Axelrad, H., Luski, I., & Malul, M. (2017). Reservation Wages and the Unemployment of Older Workers. Journal of Labor Research, 38(2), 206–227.
- Bell, A., & Jones, K. (2018). The hierarchical age-period-cohort model: Why does it find the results that it finds? *Quality & Quantity*, 52(2), 783.
- Bell, David N.F. & Blanchflower, D.G. (2011). Youth underemployment in the UK in the Great Recession. *National Institute Economic Review* 215(1), R1–R11.
- Bönke, T., Corneo, G., & Lüthen, H. (2015). Lifetime Earnings Inequality in Germany. *Journal of Labor Economics*, 33(1), 171–208.
- Chinhui, J. & McCue, K. (2017). Specialization Then and Now: Marriage, Children, and the Gender Earnings Gap across Cohorts. *Journal of Economic Perspectives*, 31(1), 183–204.
- Deaton, A., & Paxson, C. 1994. "Savings, growth and aging in Taiwan". *NBER working paper* n. 4330, National Bureau of Economic Research.
- Dietrich, H. & Möller, J. (2016). Youth unemployment in Europe business cycle and institutional effects. *International Economics and Economic Policy*, 13(1), 5–25.
- Easterlin, R.A. (1978). The economics and sociology of fertility: a synthesis. In Historical Studies of Changing Fertility, ed. C. Tilly, Princeton: Princeton University.
- Easterlin, R.A. (1987). Birth and fortune. (2nd ed.)Chicago: University of Chicago Press.
- Fannon, Z., & Nielsen, B. (2019). Age-period-cohort models. Oxford Research Encyclopedia, Economics and Finance.
- Fu, W. (2018). A Practical Guide To Age-Period-Cohort Analysis. Chapman & Hall/Crc
- Fukuda, K. (2006). Age-period-cohort decomposition of aggregate data: An application to US and Japanese household saving rates. *Journal of Applied Econometrics*, *21*(7), 981–998.
- Glenn, N. D. (2005). Cohort analysis (2nd ed.). Thousand Oaks, CA: Sage.

Graybill, F. A. (1976). Theory and Applications of the Linear Model. Belmont, C.A: Wadsworth.

- Greene, W. H. (2018). Econometric Analysis (8th Edition ed.). NY: Pearson.
- Heckman, J., & Robb, R. (1985). Using Longitudinal Data to Estimate Age, Period, and Cohort Effects in Earnings Equations. In W. M. Mason & S. E. Fienberg (Eds.), *Cohort Analysis in Social Research* (pp. 137–150). New York: SpringerVerlag.
- International Labour Organization (ILO), 2017, *Global Employment Trends for Youth 2017: Paths to a better working future*, available at: https://www.ilo.org/wcmsp5/groups/public/—dgreports/—dcomm/—publ/documents/publication/wcms_598669.pdf
- Kalwij, A. S., & Alessie, R. (2007). Permanent and transitory wages of British men, 1975–2001: Year, age and cohort effects. *Journal of Applied Econometrics*, 22(6), 1063–1093.
- Kelly, E., McGuinness, S., & O'Connell, P. J. (2012). Transitions to long-term unemployment risk among young people: Evidence from Ireland. *Journal of Youth Studies*, 15(6), 780–801.
- Kupper, L. L., Janis, J. M., Salama, I. A., Yoshizawa, C. N., & Greenberg, B. G. (1983). Age-periodcohort analysis: An illustration of the problems in assessing interactions in one observation per cell data. *Communications in Statistics - Theory and Methods*, 12(2), 779–807.

- Lipovetsky, S. 2012 "Interpretation of Shapley Value Regression Coefficients as Approximation for Coefficients Derived by Elasticity Criterion," in Proceedings of the Joint Statistical Meeting of the American Statistical Association, July–August, San Diego, CA, pp. 3302–3307.
- Lipovetsky, S., & Conklin, W. M. (2015). Predictor relative importance and matching regression parameters. *Journal of Applied Statistics*, 42(5), 1017–1031.
- McKenzie, D. J. (2006). Disentangling Age, Cohort and Time Effects in the Additive Model. *Oxford Bulletin of Economics and Statistics*, 68(4), 4.
- Mehta, A., Rysman, M., & Simcoe, T. (2010). Identifying the age profile of patent citations: New estimates of knowledge diffusion. *Journal of Applied Econometrics*, 25(7), 7.
- Nakamura, T. (1986). Bayesian cohort models for general cohort table analyses. Annals of the Institute of Statistical Mathematics, 38(2), 353–370.
- Parker, J. A. (1999). Spendthrift in America? On Two Decades of Decline in the U.S. Saving Rate, NBER Macroeconomics Annual, The University of Chicago Press, 317–370.
- Robertson, C. S. Gandini, & Boyle, P. (1999). Age-period-cohort models: a comparative study of available methodologies. *Journal of Clinical Epidemiology*, 52(6), 569–83.
- Schulhofer-Wohl, S. (2018). The Age-Time-Cohort Problem and the Identification of Structural Parameters in Life-Cycle Models. *Quantitative Economics*, 9(2), 643–658.
- Sedláček, P., & Sterk, V. (2017). The Growth Potential of Startups over the Business Cycle. *American Economic Review*, 107(10), 3182-3210.
- Shapley, L. S. (1956). Equilibrium Points in Games With Vector Payoff". US Air Force Project Rand, Research Memorandum RM-1818: 1–7
- Shimer, R. (1998). Why is the U.S. Unemployment Rate So Much Lower? In B. Bernanke & J. Rotemberg (Eds.), *NBER Macroeconomics Annual* (Vol. 13). Cambridge, MA: MIT Press, 11-74.
- Smith, C. L. (2011). Polarization, Immigration, Education: What's Behind the Dramatic Decline in Youth Employment? In *Finance and Economics Discussion Series* (pp. 41). Washington, DC: Federal Reserve Board.
- Smith, C. L. (2012). The Impact of Low-Skilled Immigration on the Youth Labor Market. *Journal* of Labor Economics, 30(1), 55–89.
- Strumbelj, E., & Kononenko, I. (2010). An Efficient Explanation of Individual Classifications Using Game Theory. *Journal of Machine Learning Research*, *11*, 1–18.
- Yang, Y., & Land, K. C. (2008). Age-period-cohort analysis of repeated cross-section surveys: Fixed or random effects? Sociological Methods & Research, 36, 297–326
- Yang, Y., & Land, K. C. (2013). Age-period-cohort analysis: New models, methods, and empirical applications. CRC Press, Boca Raton, FL.