Humanities and Social Sciences Communications

Article in Press

https://doi.org/10.1057/s41599-026-08097-9

Bridging ex-post and ex-ante EU social indicator trends: an explorative modelling analysis for SDG4 and SDG5

Received: 4 September 2025

Accepted: 20 June 2026

Published online: 11 July 2026

Cite this article as: Garcés, I., Philippidis, G., Sanjuán, A.I. *et al.* Bridging ex-post and ex-ante EU social indicator trends: an explorative modelling analysis for SDG4 and SDG5. *Humanit Soc Sci Commun* (2026). https://doi.org/10.1057/s41599-026-08097-9

Irene Garcés, George Philippidis, Ana I. Sanjuán & Rodrigo Xavier Álvarez

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

ARTICLE IN PRESS

© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

ARTICLE IN PRESS

Bridging ex-post and ex-ante EU social indicator trends: An explorative modelling analysis for SDG4 and SDG5

ABSTRACT

The 2030 Agenda for Sustainable Development represents a roadmap for navigating a sustainable path for global human development **through** 17 goals **spanning economic**, soc**ial** and environment**al domains. While ex-ante modelling is frequently used to anticipate progress in the economic and environmental spheres,** the social domain remains relatively neglected. To fill this gap and taking an EU perspective, this paper combines novel ex-post econometrics and ex-ante simulating modelling. **Using a double-lasso econometric panel specification** for two key social indicators (education and gender), **the study** identif**ies** key drivers and estimate**s** their associated effects. Results **indicate** that economic growth, lower agricultural employment, higher education expenditure and trade openness, raise post-secondary education attainment rates amongst 25–29-year-olds. **Conversely, the gender wage gap is reduced by** economic growth, increases in agricultural employment and average years of education, **more equitable income distribution** and fewer urban agglomerations. **When these drivers are integrated** into an ex-ante multi-regional simulation model, a 'global slowdown' scenario reveals a **rising** gender pay gap and **declining** post-secondary attainment rates across **the** EU **by 2040**. From a policy perspective, the drivers of tertiary attainment reinforce the importance of EU strategies that promote economic resilience, sustained public investment in human capital, and global market **integration**. **Furthermore, while** lifelong learning remains key **for narrowing the gender wage gap**, the importance of income distribution, urban agglomeration and employment in agriculture **highlights** the need for targeted labour-market and social policies **that integrate** gender mainstreaming in urban planning and rural development.

KEYWORDS: post-secondary educational attainment, gender wage gap, European Union, double-lasso, projections, simulation modelling.

ARTICLE IN PRESS

ARTICLE IN PRESS

1. Introduction and motivation

As the world's population expands and becomes more mobile, interconnected and wealthy, global human development is following an unsustainable pathway that threatens to transgress our planetary boundaries (Rockström et al., 2009). **This trajectory** carr**ies** irreversible environmental, social and economic consequences for the most vulnerable **populations**. In recognition of this existential threat, the United Nations Declaration (United Nations, 2000) laid the foundations for a sustainable route of development, through the lens of eight Millennium Development Goals (MDG) to be achieved by 2015. This **framework** was succeeded by the '2030 Agenda' (United Nations, 2015), covering a broader array of economic, environmental and social targets within 17 Sustainable Development Goals (SDGs). The creation of this consensual and comprehensive list of interdependent indicators for the economic- (e.g., growth, employment, innovation, infrastructure), social- (e.g., poverty, nutrition, health, education, gender equality) and environmental- (e.g., climate, waste, biodiversity, natural resource use) domains **serves as a vital** benchmark. **This structure enables** policymakers, academics, businesses and society **to** monitor, report and understand key indicator trends, exploring causality, and the effectiveness of policy interventions.

The European Commission (EC) has **reaffirmed** its commitment to the 2030 Agenda **with its** statistical service, Eurostat, tasked with regular monitoring and reporting on a comprehensive *ex-post* time series of SDG Indicators that reflect European values and society (Gebers et al., 2024). Unfortunately, ex-post indicator trends only yield limited insight on future developments. Global **progress** faces uncertainty arising from social, economic, environmental and technological drivers, **as well as** unforeseen high impact events or 'Black Swans' (Benton, 2019). **These extreme events** such as the financial crisis, the global pandemic, or geopolitical conflicts, can drastically reshape the path of human development, particularly within the social domain (UNESCO GEM Report, 2023; ILO, 2024).

Given this uncertainty, scenario analysis employing ex-ante simulation modelling exercises, has gathered pace (Acs et al., 2019). These research tools not only indicate the relative influence of different drivers (i.e., demographic, economic, policy, technological etc.) on indicator trends but also steer policy thinking by identifying synergies and trade-off effects **that define policy (in)coherence. Currently,** the availability of numerous ex-ante economic and environmental simulation modelling frameworks to support economic- and environmental domain indicators is relatively advanced. **These have been** even further consolidated with cross-disciplinary model linkage exercises (e.g. Hasegawa et al., 2018; Philippidis et al., 2024). Indeed, **recent** SDG focused foresight exercises **utilize** integrated assessment model approaches (e.g., Moyer and Bohl, 2019; Moyer and Hedden, 2020; Moreno et al., 2024**). These approaches link specialised models covering different fields (e.g.,** health, demographics, economics**,** energy) to a harmonised baseline, each **contributin**g their own **specific** list of SDG indicators. Despite the computational expense in aligning so many models,

ARTICLE IN PRESS

these exercises undoubtedly provide deep insights into synergies and trade-offs between SDGs. Notwithstanding, Orbons et al., (2024) reveal from a meta-analysis of 140 papers, that the coverage of social indicators, particularly SDG5 (gender equality), remains relatively sparse. This **deficit is attributed to** the difficulty in quantifying the "social and political processes that dictate the level of achievement of these goals" (pp3, Orbons et al., 2024). **Ultimately**, this **underscores** the general weakness that ex-ante modelling is ill equipped to endogenously capture the complexity of 'institutional' change, which is a key influencer of social indicators.

A computationally less expensive alternative for capturing social indicators is based on the use of **an ex- ante** computable general equilibrium (CGE) macroeconomic model fused with **an ex-post** panel econometric approach **(Campagnolo et al., 2018; Campagnolo and Davide, 2019; Moreno et el., 2024). By accounting for the behaviours of numerous economic agents (i.e. firms, households, governments, and investors), the CGE framework provides a rich variety of proxy variables that reflect institutional progress. These proxy variables can then be linked to statistically estimated behavioural parameters to reflect shifts in social indicators**. A glance at these studies reveals examples of social indicators in the domains of poverty (SDG1) (Campagnolo and Davide, 2019), nutrition (SDG2) (Moreno et al., 2024), health (SDG3) (Campagnolo et al., 2018; Moreno et al., 2024), education (SDG4) (Moreno et al., 2024) and income inequality (SDG10) (Campagnolo and Davide, 2019). **However,** indicators for gender equality (SDG5) are not present. **Furthermore,** the published studies **lack a** justification of the econometric model specification that provide **their** parameter estimation.

Taking an EU focus, this paper has four principle aims that highlight the contribution to the literature. Thus, the paper seeks to (i**) address** the current paucity of foresight simulation studies dealing with SDG social indicator outcomes by employing an **integrated** ex-post/ex-ante bridging technique; (ii) improve on ex-post model specification criteria by **implementing a rigorous** data selection and reduction procedure based on machine learning. **Specifically,** a double least absolute shrinkage and selection operator (lasso) model is employed to identify robust drivers of social indicators (see Belloni et al. 2014; Chernozhukov et al., 2018); (iii) **fill** a gap in the foresight literature by applying the ex-ante/ex-post bridging approach to an SDG5 indicator (gender wage gap computed as the difference in mean gross hourly earnings between male and female employees, expressed as a proportion of mean male earnings from the Eurostat database); (iv) compare the econometric outcomes of our model with prior research (i.e., model structure and estimates) for a specific education (SDG4) indicator (post-secondary educational attainment rate among individuals aged between 25–29 sourced from the Wittgenstein Centre). **To achieve these aims, the study pursues three supporting objectives**, where we seek to (a) conduct a literature review for each social indicator to develop a comprehensive panel database of candidate drivers for data selection and estimation (b) modify the CGE model code and relevant data to calibrate and baseline said social indicators and (c) explore forward-looking system-wide uncertainty by

ARTICLE IN PRESS

examining how these drivers shape the trajectory of both social indicators under an alternative 'global slowdown' scenario.

Our main research hypotheses are focused on the expected sign and magnitude of the social indicator drivers (see related literature, section 2). **These drivers determine** the relative impacts on each social indicator in the CGE model when compared with the CGE baseline. Accordingly we formulate the following four hypotheses: (i) **Gross Domestic Product (**GDP) per capita growth is associated with rising post-secondary attainment due to increased household resources and public education investment, while it decreases the gender wage gap; (ii) structural transformation (e.g. lower employment in agriculture and greater trade openness) is associated with higher post-secondary educational attainment; (iii) **an increase in the average** years of schooling are associated with a lower gender wage gap as education enables women to access more skilled and better-paid occupations; (iv) greater income inequality reflecting unequal access to opportunities and labour market segmentation increases the gender wage gap.

The remaining sections of the paper are organised as follows. Section 2 presents the literature review of ex-post empirical studies of candidate drivers of education and gender equality. Section 3 describes the ex-post econometric modelling stage, including the data used, the lasso technique and the results. Section 4 presents the ex-ante CGE model focusing on the implementation of the econometric estimates and results of the simulations under uncertainty. Section 5 concludes.

ARTICLE IN PRESS

2. Related literature

The right of access to education and gender equality is enshrined within the 1948 United Nations Declaration of Human rights. However, the OECD (2025) report still documents persistent inequalities in access to tertiary education. Improved access to education enhances the human capital stock, thereby raising productivity and economic growth, whilst promoting equality and poverty reduction (Barro, 2013). **Beyond economic metrics,** raised levels of education also have positive impacts on issues ranging from health, fertility, family structure, political participation, and crime reduction (Craigwell et al., 2012; Urhie et al., 2014). Likewise, **the** reduc**tion in** gender inequality **is a catalyst for** employment (European Institute for Gender Equality, 2017), real GDP growth (Dabla-Norris and Kochhar, 2019) and financial stability and economic diversification (Fernández et al., 2021). Research on the determinants of educational attainment and gender inequality has expanded significantly in recent years, yet important gaps remain. Recent studies emphasise the role of structural economic forces in shaping tertiary education **w**hile the literature on gender wage gaps has increasingly focused on institutional and structural mechanisms.

In seeking to estimate the strength of the drivers of educational attainment and the gender wage gap, there are examples of standard regression analysis (i.e. ordinary least squares-OLS or two stage least squares, Dridi, 2014) or, in the case of gender equality, micro decompositions (i.e.

ARTICLE IN PRESS

Oxaca-type, Kunze, 2018). **However, these approaches often pay** limited attention to high dimensional settings, macro structural drivers or model uncertainty (Blau and Kahn, 2017; Athey and Imbens, 2019). **Among these drivers,** the change in real GDP per capita is widely recognised as a key driver as it reflects overall spending power (Dridi, 2014) and boosts the demand for 'normal' goods such as educational services (Craigwell et al., 2012). **In contrast,** the relationship between public spending and educational attainment, **remains ambiguous** (Craigwell et al., 2012, Senadza and Hodey, 2015). **Empirical evidence offers** mixed results **which** may partly be explained by the 'crowding out' of private demands by public spending or by inefficiencies linked to institutional limitations and weak budgetary management (Rajkumar and Swaroop, 2008). **Furthermore,** the OECD (2023) highlights widening cross-country disparities linked to public investment, demographic pressures and regional inequalities.

A further strand of literature focuses on the impact of trade openness on the demand for higher education. Autor et al., (2013) **demonstrate** a positive relationship exists in regions more exposed to import competition or that have undergone a trade liberalisation expanding the demand for more highly skilled and educated workers. Similarly, Blanchard and Willman (2016) show that freer trade **incentivises** high skill acquisitions **while** induc**ing the** crowding out of the middle occupations. Other studies **identify** an inverse relationship between agricultural employment and educational attainment (Caselli and Coleman II, 2001; Herrendorf and Schoellman, 2018). As economies develop and diversify, **rising education levels typically facilitate a structural** shift away from agriculture into secondary and tertiary sectors **which demand** specific **technical** skills and qualifications. **Consequently, t**hose with post-secondary or higher education are more likely to be employed in non-agricultural sectors that offer higher wages and more specialised job opportunities.

Regarding gender equality, the literature attributes the gender wage gap to supply or demand factors. **Establishing c**ausality from supply factors **is challenging** due to endogeneity issues. **For instance, the** non-random selection of women into **the labour force** introduces **estimation** bias **as** the decision to return to work after childbirth as well as **subsequent** interruptions are endogenous (Blau et al., 2024).[i] **Furthermore, w**orking part-time is also likely to negatively affect career progression (Kunze, 2018) and leads to downgrading in terms of occupations (Manning and Petrongolo, 2008).[ii] Goldin (2014**) observes** that while gender pay gap differences are **minimal at the start of the career**, they **widen** with age. **This is particularly evident** in high-skilled professional occupations **where** men are more likely than women to work long hours. **In certain sectors**, these hours are disproportionately rewarded, a situation referred to as a nonlinear wage-hours schedule. **Conversely, the increasing educational attainment of women relative to** men has contributed to a decrease in the gender wage gap. Similarly, years of work experience have also contributed to **reductions** in the gender wage gap even more intensely than years of schooling (Blau and Kahn, 2017). On the demand-side, **factors contributing to the wage gap have been found to result from employer prejudices. For instance,** Correll et al. (2007) find that in a recruitment process in the United States, mothers were

ARTICLE IN PRESS

perceived less favourably than non-mothers, even though there was no statistical difference for fathers *versus* non-fathers. While Benson et al., (2026) show that subjective assessments of employee "potential" also contribute to gender gaps in promotion, pay and to an inefficient allocation of talent across roles.

Other structural factors **intersecting** with supply and demand factors suggest that job mobility (Kunze, 2018) and psychological attributes **such as** non-cognitive skills (Blau and Kahn, 2017) can lower initial career aspirations, contributing to the gender wage gap (Azmat and Ferrer, 2017). Elsewhere, another perceived driver is the 'relocation' of gender wage gap inequality (Quadlin et al., 2023). Wh**ile** the largest wage gaps were once concentrated among lower-paid, lower-educated workers, these wage gaps **are now most prominent among the** highest-paid, highly-educated workers where top **tier** earnings **growth** have far outpaced **gains at the** lower **end of** the wage distribution (Piketty et al., 2018; Haltiwanger et al., 2024). This suggests that economic returns for a college degree among the top 10 percent earners have grown faster for men than for women (Mandel and Rotman, 2021).

Likewise, urbanisation and agglomeration have been posited as **influential factors**. **Specifically, while urban centres offer** access **to specialised** jobs **un**available in non-urban areas (Nisic, 2017), long distances and commuting lead women to accept lower wages i**n exchange** for shorter commutes (Liu and Su, 2024) or **transition to** part-time jobs (Bartoloni et al., 2023). Moreover, evidence suggests that **sectoral composition is** a **key** factor. While lower wages for women **are still** reported in agricultural sectors, even across member states of the EU where women account for around 30% of farm ownership (Diamanti and Duncan, 2023), there appears to be less wage dispersion. **In contrast, the gender wage gap is more pronounced in high growth industries such as** financ**e** or IT sectors (Leythienne and Pérez-Julián, 2021).

In both settings, these drivers are subsequently inserted as behavioural parameters into the ex-ante simulation model to explor**e** forward-looking system uncertainties. Although ex-ante modelling of social indicators remains limited, recent contributions point to growing interest in integrating social dimensions into long-term scenario frameworks. The European Commission's Joint Research Centre Work Program (2025-2027) stresses the importance of scientific evidence, analytical tools and anticipatory approaches to support EU policy making. **This strategic document reinforces** the relevance of more data driven approaches to parameter estimation. **Consequently, our integrated** ex-post/ex-ante framework provides a novel methodological contribution to the social indicator modelling literature and a tool to anticipate policy effects.

3. Econometric modelling and results

3.1. Data sources

An EU region focused panel data set is constructed for a large selection of candidate

ARTICLE IN PRESS

explanatory variables or drivers of both post-secondary educational attainment rates and the gender wage gap, based on the literature review and *a priori* expectations.

3.1.1. Educational attainment rates

The final sample to analyse the drivers of post-secondary educational attainment rates covers the period 1995-2020 and includes the current 27 EU members, plus the United Kingdom, Norway and Switzerland. With global coverage, the main data sources are the Wittgenstein Centre for Demography and Global Human Capital (see Lutz et al., 2018, KC et al., 2024) and the World Development Indicators database (World Bank, 2025). Wittgenstein provides observed data up to 2020 **which** is a relevant point to end the sample to avoid noise introduced by exogenous shocks such as COVID or the Ukraine war. Indeed, UNESCO's Global Education Monitoring Report (2023) highlights that economic shocks such as the COVID19 pandemic and subsequent inflationary pressures have disproportionally affected young adults' education performance. **Similarly,** national-level research confirms that the COVID19 pandemic has adversely affected education performance in many EU countries (European Commission's Education and Training Monitor, 2024).

Thus, from Lutz et al. (2018), we obtain observations on post-secondary educational attainment rates amongst 25–29 year-olds. The data reveals that on average, post-secondary education attainment rates have risen from less than 30 to 44 percent, with considerable heterogeneity across regions, both in the initial value and growth rates (See Fig. A1 in the online resource A.2). **Beyond the main education indicator,** this same data source **provides** time series observations **for** the gender gap in years of schooling, fertility rates and changes in population size.

From the World Bank (World Development Indicators database), economic development data is used (i.e., per capita GDP in constant 2015 US$), income distribution measures (the lowest and highest 10 percent of income share), labour force participation (i.e. total labour force, employment in agriculture as a percentage of total employment), population data (i.e., the percentage of the population in urban agglomerations of more than 1 million of the total population, the percentage of the population in the largest city), education indicators (i.e. the percentage of children out of school of primary school age and government expenditure on education as a share of GDP) and trade data (i.e. trade openness defined as a share of the sum of imports and exports divided by total GDP in constant 2015 US$). In the online resource A.2, Table A2 further describes these variables while Table A3 provides the summary statistics. In addition to these ex-post time series observations, projected values of post-secondary educational attainment rates to 2040 are taken from KC et al. (2024), which are used in the ex-ante CGE analysis (see later). This required smoothing of the connection between the ex-post data that ends in 2020 (Lutz et al., 2018) and ex-ante projections that start one year later until 2040.[iii]

ARTICLE IN PRESS

3.1.2 Gender Pay Gap

Defined as the ratio of the difference between average gross hourly earnings of male and female employees, to the average gross hourly earnings of male employees, observations of the gender pay gap (dependent variable) come from Eurostat SDG 5 (gender equality). The data starts in 2002 and consistently with the education attainment variable, finishes in 2020 and covers all EU member states, the United Kingdom, Norway and Switzerland. In 2020, **the average gender wage gap** for the EU27 **stood at** 12.9 indicat**ing** that on average, men earn 13% more than women, per hour. **However, this aggregate** reveal considerable differences **at the member state level**, ranging from below 5% in Luxembourg, Romania, Slovenia, Poland, Belgium and Italy to more than 17% in Hungary, Germany, Austria, Estonia and Latvia. In most countries, the gender wage gap has decreased over the last twenty years (see the online resource A.2, Fig. A2).

For consistency with the dependent variable, data for drivers such as real GDP per capita, the proportion of women in parliament, the corruption perception index and the skilled employment share for males and females, are also retrieved from Eurostat. These data are supplemented by variables from Lutz et al. (2018) on fertility rates, the crude birth rate which measures the number of live births occurring per 1,000 people and the population size of the country. Moreover, from the World Bank (2025), data on income distribution, labour force, population and trade measures similar to the ones used for educational attainment are incorporated. In online resource A.2, Table A4 presents more details on the description and definition of the variables while Table A5 summarises the descriptive statistics of the variables considered in the model for gender.

ARTICLE IN PRESS

3.2. Estimation

Given the multitude of potential drivers for post-secondary educational attainment rates and the gender wage gap, a key aim of the current research is to objectively prioritise their statistical importance. For this purpose, the lasso technique is applied which helps in selecting a subset of important predictors from a large set of variables, thereby reducing the complexity of the model whilst limiting the role of the researcher in deciding which configuration of variables to include in the final model. As the failure to control for valid covariates can lead to biased parameter estimates (Urminsky et al., 2016), it is important to start with **the** broadest available set of variables. In the application of lasso, a similar approach to Adeleke and McSharry (2022) is followed, although in this paper, a double-lasso technique (Belloni et al. 2014, Chernozhukov et al., 2018) is applied (see below).

Lasso is a regression technique that combines variable selection and regularisation to improve the predictive accuracy and interpretability of the final statistical model. Lasso adds a penalty or regularisation, which shrinks the estimated coefficients towards zero resulting in some coefficients being exactly zero and consequently, excluding them from the model. The shrinkage property is

ARTICLE IN PRESS

beneficial when the number of observations is small relative to the number of predictors (James et al., 2013). However, directly using the lasso regression may under-estimate those variables that are non-zero leading to their exclusion even in cases where they are relevant. The double-lasso approach solves this problem by identifying covariates that predict **not only** the dependent variable but also **the** focal variable, which the researcher considers of high relevance.

In short, we apply the double-lasso as follows:

Step 1: estimates a lasso regression predicting the dependent variable (Y_{it}) and saving explanatory variables (W_{it}^{j} with $j = 1, \dots, K$) with non-zero estimated coefficients (α_j):

$$Y_{it} = \alpha_0 + \alpha_1 W_{it}^{1} + \cdots + \alpha_K W_{it}^{K} + \varphi Z_{it} + \varepsilon_{it} \quad (2)$$

where Z_{it} represents the focal variable and subscripts *i* and *t* refer to country and year, respectively.

Step 2: fits a lasso regression predicting the focal independent variable (Z_{it}) and keeps track of the variables with non-zero estimated coefficients:

$$Z_{it} = \delta_0 + \delta_0 W_{it}^{1} + \cdots + \delta_K W_{it}^{K} + \varepsilon_{it} \quad (3)$$

As a focal independent variable, we choose GDP per capita as it has been proven to be relevant in the literature (Urhie et al., 2014; Jayachandran, 2015).

Step 3: fits a linear regression of the dependent variable on the focal independent variable, including the covariates W_{it}^{k} selected in either of the first two steps. This is relevant because the exclusion of a covariate that is a modest predictor of the dependent variable, but a strong predictor of the independent focal variable can create a substantial omitted variable bias (Darlington, 1990):[iv]

$$Y_{it} = \beta_0 + \beta_1 Z_{it} + \sum_{1}^{K} \beta_k W_{it}^{k} + \varepsilon_{it} \quad (4)$$

The double-lasso technique helps in exploiting the cross-sectional variability and the likely low temporal variation in our dataset while country heterogeneity is largely captured by **a** rich array of control variables. Moreover, double-lasso particularly helps at mitigating endogeneity coming from omitted variable bias.[v]

Some final practical issues in the double-lasso **implementation** are the following. Firstly, amongst alternative options on the selection of the penalty parameter λ, we use the plug in method that has been shown to perform well in a multitude of settings **that** neither under-**penalises** nor over-**selects** variables (Hirukawa et al, 2023). Secondly, the set of possible explanatory variables considered in the first step leads to a final model with a low mean squared error (MSE) which generally indicates a better model fit. Finally, most of the explanatory variables used in both, education and gender regressions, are ratios or percentages, and a minority are continuous. To minimize differences in scale and **enhance estimation** efficien**cy**, both explanatory and dependent variables are transformed into logarithms (Gerdes, 2011). Furthermore, **this log-log**

ARTICLE IN PRESS

specification allows the interpretation of coefficients directly as elasticities, which is **also** convenient for insertion into the CGE model.

3.3. Econometric Results

Table 1 presents the final outcome of the double-lasso **regression** for both dependent indicators, education attainment and gender wage gap. As a reminder, the full set of covariates considered is presented in Tables A2 and A4 for education and the gender wage gap, respectively. A total of 8 and 9 of these covariates are retained as controls in the final regression shown in Table 1. The final set of explanatory variables in Table 1 is selected by lasso based on joint significance (Wald test p=0.000 in both models). **Detailed findings for** each indicator are discussed in the following sub-sections.

Table 1. Double-lasso results.

	Post-secondary education attainment	Gender wage-gap
GDP per capita	0.143^{***} (0.035)	-0.358^{***} (0.091)
Employment agriculture *(% total employment)*	-0.121^{***} (0.047)	-0.384^{***} (0.072)
Education expenditure *(% GDP)*	0.168^{**} (0.077)	-
Trade openness $((Imports + Exports)/GDP)$	0.179^{***} (0.059)	-
10 percent highest income *(% of income)*	-	1.175^{***} (0.408)
10 percent lowest income *(% of income)*	-	0.983^{***} (0.148)
Urban agglomeration *(% population in cities of more than 1 million over total population)*	-	1.352^{***} (0.183)
Corruption perception of judiciary system (scale from 0 to 100)	-	0.364 (0.229)
Mean years of schooling 20-24	-	-2.085^{**} (0.853)
N	278	163

Note: All variables are in logarithms as recommended by Gerdes (2011). Standard errors in parentheses: $^{*}\ p < .10$, $^{**}\ p < .05$, $^{***}\ p < .01$. For more details on the full set of variables included in the first step from which lasso selects the ones shown in this table please see Table A2 and Table A4 in online resource A.2.

3.3.1. Drivers of post-secondary educational attainment

The first column in Table 1 summarises estimation results for education drivers. The indicators explain**ing** most of the variation in postsecondary education attainment relate to structural factors that influence household resources, public investment capacity and labour market demand for high skilled workers such as GDP per capita, education expenditure, trade openness. Consistent with **existing** literature (Rajkumar and Swaroop, 2008; Senadza and Hodey, 2015), the results show

ARTICLE IN PRESS

that GDP per capita is a statistically significant driver, with a 1% higher per capita GDP leading to a 0.143% increase in post-secondary educational attainment.

The estimation results further highlight the role of structural transformation and public investment. A **1%** rise in the employment share of agriculture is shown to statistically reduce the level of post-secondary educational attainment by 0.121%. This **result aligns with** the inverse relationship of development and agricultural employment (Caselli and Coleman II, 2001; Herrendorf and Schoellman, 2018), **as** individuals with higher education levels **typically gravitate towards** non-agricultural sectors offer**ing** wage **premiums** and specialised job opportunities. **Regarding public policy, a 1% rise in the education** expenditure share is statistically associated with a 0.168% increase in post-secondary education attainment rates**, a finding** consistent with previous literature (Senadza and Hodey, 2015). **When compared to** recent **findings** by Moreno et al. (2024), the **direction** of influence **coincides though our reported** magnitude is smaller. This **discrepancy** may be attributed **to our methodological differences: we employ a double-lasso technique and a broader set of controls, whereas their study relies on a** country-specific panel **and** uses urban population shar**e as the sole control variable.**

The **empirical evidence presented here provides a robust validation of the trends highlighted in the** European Commission's Education and Training Monitor (2025) report. **Specifically,** tertiary education attainment continues to rise across the EU **though significant disparities remain largely dictated by** GDP per capita and urbanisation levels, with rural areas reaching a lower achievement. Our **findings** on the positive **influence** of GDP per capita and urbanisation **align** with these EU results, while the negative impact observed for employment in agriculture can be viewed as coherent with the EU result on the lower attainment in rural areas. **We also find that public investment is a critical lever. For example, several** European countries provide direct and indirect financial support to **boost** higher rates of tertiary education, showing that higher public expending can increase program completion (European Commission's Education and Training Monitor, 2024). Finally, trade openness has a positive and significant effect on post-secondary education attainment, standing out as the driver with the highest elasticity (0.179). This positive impact **supports existing** literature **suggesting that global trade can be a t**rigger for the demand of high skilled and educated workers (Autor et al., 2013).

3.3.2 Drivers of the gender wage gap

The second column in Table 1 summarises the estimation results for the gender wage gap drivers. **Our results show that the** gender wage gap is influenced by more complex institutional and labour market structures such as labour market dynamics, spatial effects and inequality **as suggested by** OECD, (2025). This heterogeneity naturally produces stronger effects for some variables (e.g., mean years of schooling, income distribution, urban agglomeration) and weaker or more context dependent effects for others (e.g. employment in agriculture and GDP).

ARTICLE IN PRESS

The negative coefficient for GDP per capita (-0.358) highlights that increases in the level of economic development are statistically associated with **a contraction** in the gender wage gap. This **trend likely reflects the** positive effect of gender equality policies in more developed countries. **Results further indicate that a 1%** rise in the agricultural employment share decreases the gender wage gap by 0.384%. **This outcome can be attributed to the** lower income **levels typical of the agricultural sector** compared to the manufacturing or services sectors (Schuh et al., 2019), **which as we observed in the literature review section, results** in less wage dispersion **relative to higher growth industries** (Leythienne and Pérez-Julián, 2021).

Regarding income distribution, the higher the proportion of income held by the richest ten percent of the population, the greater is the gender wage gap, a result that is consistent with Quadlin et al. (2023) and England et al. (2020). **This might be attributed to the disproportionate income growth of men in high-level executive positions.** Interestingly, the results also reveal that when the share of income held by the poorest ten percent increases, the gender wage gap also increases. **This may be** due to an increasing number of women working in particularly low paid and/or temporary jobs.

Supporting the findings of Liu and Su (2024) and Bartoloni et al. (2023), a higher proportion of the population in urban agglomerations of more than one million individuals increases the gender wage gap. **This phenomenon may be driven by specific labour market trade-offs. For instance**, women **in large metropolitan areas** accepting part-time jobs or lower paid jobs with smaller commutes (Liu and Su, 2024). Urban areas can generate opportunities related to income, employment, food security, access to education and health services, social protection, and communication and transport. However, agglomerations pose potential threats that may arise in the social sphere like an increase of inequality, insecurity, violence, congestion of infrastructures or environmental problems.

The perception of corruption, which measures respondents' perceptions regarding the independence of the judiciary system in a country, was selected as relevant by lasso, but was not found to be significant. **Lastly, as the variable with** the strongest absolute impact on the gender wage gap, a one percent rise in mean years of schooling is found to reduce the gender wage gap by 2.085%. This is consistent with the **study** of Blau and Kahn (2017).

Finally, it is important to note that EU focused studies (European Commission, 2025; OECD, 2025) document persistent gender gaps despite rising female educational attainment (which has been the largest contributor in our regression). **This points** to the role of other structural factors such as labour market segmentation, demography effects, spatial inequalities and institutional weaknesses consistent with the elasticities identified in our model that highlight stronger effects for variables regarding income distribution, urban agglomeration or employment in agriculture.

3.4. Projections of social indicators for ex-ante analysis

ARTICLE IN PRESS

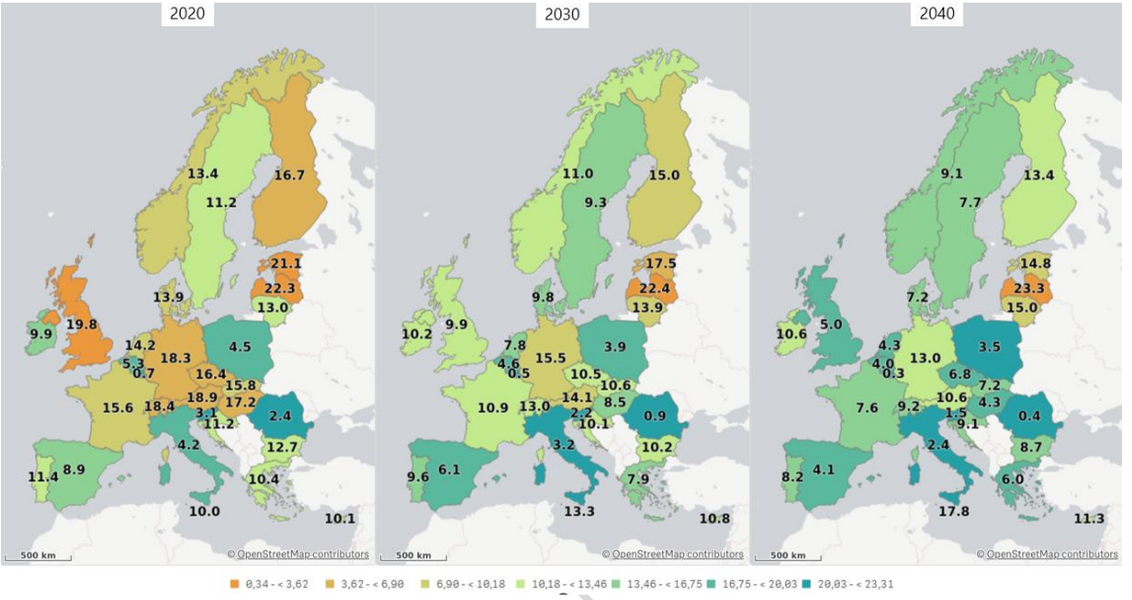
For post-secondary educational attainment rates amongst 25–29-year-olds, projections are s**ourced** from KC et al., (2024) **representing different** global change scenarios. These so-called Shared Socioeconomic Pathways (SSPs) are designed to represent alternative future trajectories regarding socioeconomic challenges for both climate change mitigation and adaptation (O'Neill et al., 2017). **The most recent** SSPs provide a much richer picture of economic, demographic, and social trends as well as energy and economic variables (Riahi et al., 2017). The five SSPs are: SSP1 (Sustainability / Rapid Social Development); SSP2 (Middle-of-the-Road / Medium Population Scenario); SSP3 (Fragmentation / Stalled Social Development); SSP4 (Inequality / A Road Divided); SSP5 (Fossil-fuelled Development / Taking the Highway). We focus our analysis on SSP2 which projects **a continuation of current trends characterised by moderate** progress toward achieving development goals, reductions in resource and energy intensity, and **a gradual** decrease **in** fossil fuel dependency. Fig. B1 in online resource B presents the projection **increases** of post-secondary education attainment by country. In section 4.1, these projections are introduced into the CGE model.

To the best of our knowledge, there are no multi-region long-term projections for the gender wage gap. Thus, in-house projections are constructed using the predictive capacity of our econometric model coupled with projections of its available drivers. For GDP per capita we use the projections retrieved from the different SSPs[vi] while for mean years of education for individuals between 20-24 years of age for each SSP, KC et al., (2024) data is used. For **those** explanatory variables **lacking** projections by an international **organisation** (share of income owned by the highest or lowest ten percent, corruption perception, urban agglomerations or employment in agriculture), we **estimate them** for the period 2021-2040 based on an OLS fixed effects model **with** country-specific linear trends.[vii] The final projection for the gender wage gap as well as the coefficients estimated are then introduced in**to** the CGE model (see section 4**.1**).

Figure 1 **illustrates** our gender wage gap projections by EU member states for the SSP2 scenario up to 2040. This final projection is a combination of the **projections** of each country specific driver and the parameters obtained from the econometric model. Overall, reductions in wage gap inequality are expected in most countries over the twenty-year period (Fig. 1). These reductions are motivated by increases in mean years of schooling and GDP per capita and decreases in the percentage of income held at the highest and lowest decile. Although employment in agriculture **serves to reduce** the gender wage gap**, its overall share is small and projected to decline during this period, thereby** limiting its cumulative contribution. **Conversely**, five EU member states (i.e., Cyprus, Ireland, Lithuania, Latvia, Malta) **are projected to experience** mild increases in gender wage inequality. These effects are associated with stronger increases in urban agglomeration than in other countries, as well as increases in the percentage of income held by the poorest decile (i.e., Latvia, Lithuania, Ireland) or by the highest decile (i.e., Cyprus, Lithuania and Malta).

ARTICLE IN PRESS

Fig. 1 Gender wage-gap projections by country in 2020, 2030 and 2040 SSP2 (Middle of the road)



Source: own elaboration based on the estimated model (Table 1, second column). Values for explanatory variables explained in the text. The figure shows observed values for 2020 and our own projections for 2030 and 2040.

4. CGE modelling and results

4.1. CGE model

The study **employs** the Modular Applied General Equilibrium Tool (MAGNET) simulation model (Woltjer et al., 2014**) as an advanced variant of the neoclassical multi-regional CGE**.[viii] MAGNET includes an array of non-standard specialist modules of code, including a dedicated SDG module (Philippidis et al., 2020), **that can be** activated using binary switches. With the well-known multi-region 'Global Trade Analysis Project' (GTAP) model (Corong et al., 2017) at its core, the MAGNET model structure is a closed system of simultaneous equations consisting of (i) behavioural equations for representative agents (i.e., households, producers, exporters, importers, public sector, investors, savings), (ii) market clearing equations that equate supply and demand in all markets and (iii) accounting equations that both ensure zero 'economic' profits, and a net-zero balance of payments. **The model employs the GTAP 11 database (Aguiar et al., 2023) with 160 regions and 65 activities, with a benchmark year of 2017, the latest available year**. **The database is constructed** on the principles of a square social accounting matrix (SAM) **where transaction values for** final, intermediate and investment demands (columns) equal the **total** value of their **corresponding** commodity supplies (rows). The behavioural **parameters** are subsequently 'calibrated' to the GTAP global transactions values**.**[ix]

ARTICLE IN PRESS

The macroeconomic model contains 'n' equations and 'm' variables. **To reach a unique solution, 'n' variables are defined as** endogenous (i.e., determined by the model), **while the remaining** 'm-n' variables are held exogenous (i.e., determined externally by the modeller). **Exogenous variables** are typically changes in population, primary factor endowments, technology ratios and tax/subsidy rates. By 'shocking' one or more exogenous variable(s) from its initial equilibrium point, a new set of market clearing prices and quantities is obtained at the counterfactual equilibrium point. As the CGE model is a closed economic system**, it** internalises all inter-industry competition effects for primary factors (e.g., land, labour, capital) and inputs, as well as the limits on household and public demands subject to income constraints**. T**he model **also** fully captures the resulting multiplier impacts and rebound effects arising from an initial market 'shock'.

4.1 Implementation of both social indicators into MAGNET

To demonstrate the model's application, a 'business as usual' baseline consistent with SSP2 is constructed from 2017 to 2040 with time intervals for 2020, 2025, 2030 and 2035. A regional disaggregation of EU member states is chosen, plus additional non-EU partners. The baseline is principally driven by exogenous projections of real GDP, population, labour force and capital stock changes activated by regional rates of return.[x]

As the behavioural elements of the MAGNET model operate in percentage changes, the additive log-log functional form **equations** of estimated drivers **for each social indicator** from Table 1 can be **coded** directly into **MAGNET**. The**se coefficient** estimates **serve** as response parameters (elasticities) in the CGE model, multiplied by endogenous percentage changes in their respective drivers calculated from the CGE model**.**[xi] In both log-linear equations, an additive error term is **retained to accurately target baseline projected social indicator values** (see below). **For the 2017 benchmark year of the model, starting values of both social** indicators **across** all EU member states **are inserted into MAGNET's database.**

In seeking to mimic the expected path of inputted SSP2 projection trends for both social indicators, **we address the inherent** structural incompatibilit**y between the econometric estimates and the CGE environment. While the econometric** estimat**ion assumes a ceteris paribus condition** (**i.e.** all other market forces are constant), **the CGE model assumes simultaneous changes in all 'n' markets.** To reconcile these differences, the calibration of our baseline up to 2040 consists of two steps. First**ly**, **we construct** a pre-simulation baseline **to 2040 including the** exogenous SSP2 **trajectories** to both social indicator variables for all EU member states **utilizing our econometric projections for the gender wage gap and the Wittgenstein Centre projections (KC et al., 2024) for post-secondary education (section 3.4). C**ompensatory adjustments in the endogenous additive error term in each equation guarantee that the **pre-targeted** indicator **projections** are **met by MAGNET**. Second**ly**, the same experiment is rerun

ARTICLE IN PRESS

where calibrated errors **terms** from the **first step** are exogenised and shocked by period. **This ensures that** the (now) endogenous social indicators follow their baseline trajectories. **The model is then ready for further 'scenarios', where** deviations from the baseline **trend result from endogenous marginal changes in MAGNET's drivers**. To illustrate the disruptive impacts of uncertainty, **we model** a scenario **representing** a relative slow**down in global** real GDP growth. This is modelled via a cumulative relative 5% reduction in region-wide productivity in all world regions between 2025 to 2040.

4.2 CGE social indicators results

The baseline average post-secondary educational attainment rate **for the EU** amongst 25-29-years-olds **starts at an absolute level of** 41.38%, **projected to rise** to 53.68% by 2040 (not shown). **Under** the global slowdown scenario, this **rate reaches** 51.09%, **representing a** relative fall of -4.83% change compared with the 2040 baseline (Fig.2)**. A** decomposition of the MAGNET model drivers (Fig.2, left side panel) **reveals the specific factors contributing to this result. In** order of magnitude, the relative fall in real GDP per capita, relative rise in the agricultural employment share and the falling education expenditure share, reduce relative rates compared with the baseline of post-secondary attainment by -2.82%, -2.17% and -0.35%, respectively. In contrast, **trade openness provides a** slight relative **positive contribution of 0,51% to the indicator.**[xii] **This occurs** due to quicker falls in the value of EU GDP relative to the fall in the value of the EU27 bloc's trade.

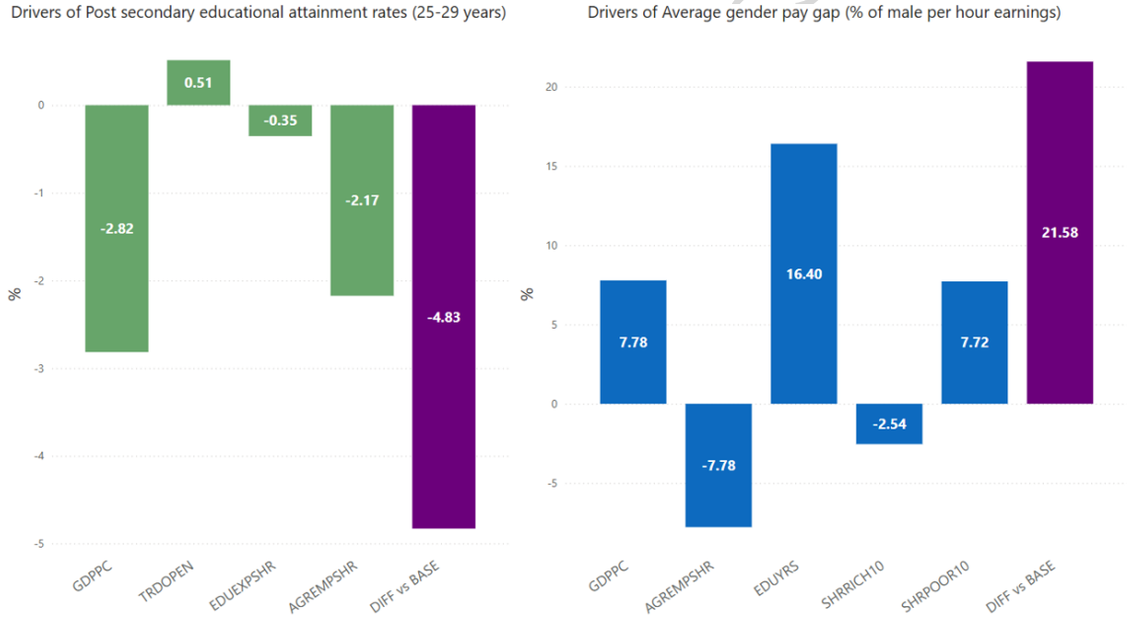
An examin**ation of** EU **countries under** the global slowdown scenario in 2040 **reveals a consistent downward trend across the block. As illustrated in** the left side panel of Fig.3, the relative falls in the education indicator compared **to** the baseline **is broadly uniform with results** closely distributed around the EU27 average.

Expressed as a proportion of the male employee gross hourly wage, in the baseline the absolute value of the gender wage gap indicator falls from 13.31% in 2017, to 6.86% in 2040 (not shown). The global slowdown scenario results in an average absolute EU-wide gender wage gap of 8.34% by 2040, equivalent to a relative rise of male gross hourly wages to female gross hourly wages of 21.58% compared with the baseline (right side panel, Fig. 2). A decomposition analysis of the MAGNET model drivers for the EU27 average **identifies the factors contributing to this widening gap. Relative to** the baseline (in order of relative magnitude), falls in the average years of education amongst 25-29 year olds and in real per capita GDP, as well as a relative rise in the concentration of wealth in the poorest 10% of the population, raise the relative gender wage gap by 16.40%, 7.78% and 7.72%, respectively. In contrast, the rising agricultural employment share and the fall in the share of wealth in the richest decile, reduce the relative gender wage gap by -7.78% and -2.54%, respectively.

At the member state level, the right-side panel of Fig. 3 reveals **significant variation in the**

relative **rise of the** gender wage gap across EU member states, as well as **in** the absolute indicator levels for the 'global slowdown' scenario in 2040. With the same general driver trends as reported for the EU average, relative rises in gender pay inequality range between 5.12% (Italy) **to** 40.42% (Belgium and Luxembourg). Whilst most of the MAGNET drivers have broadly similar impacts across the member states, the driver of average years of education introduces cross region heterogeneity. **This variance is driven by** changes in its own region-specific drivers of per capita food demand (relative falls) and the share of agriculture in value added (relative rises). The former is determined by the region-specific income elasticities of demand and the latter by the relative magnitudes of the general equilibrium supply elasticities in the agricultural and non-agricultural sectors.

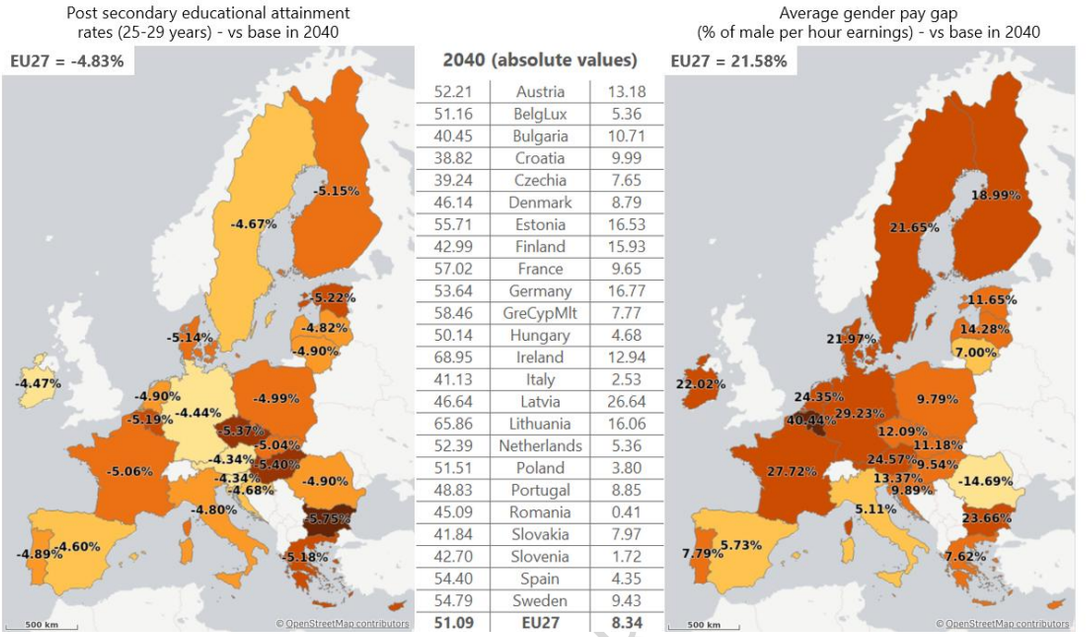
Fig. 2 Cumulative drivers of social indicators vs the baseline in 2040



Source: own elaboration based on the estimated model (Table 1) combined with projections in section 3.4 and additional data from MAGNET. Note: GDPPC is GDP per capita, TRDOPEN is trade openness, EUEXPSHR is the education expenditure share, GRMPSHR is the agricultural employment share, EDUYRS is mean years of schooling for the population between 20-24 years, SHRRICH10 and SHRPOOR10 is the concentration of income held by the richest and poorest 10 percent, respectively. Diff vs Base shows the difference in percentage terms between the baseline and the simulation scenario.

ARTICLE IN PRESS

Fig. 3 Changes in education and gender indicators vs baseline in 2040



Source: own elaboration based on the estimated model (Table 1) combined with projections in section 3.4 and additional data from MAGNET.

5. Conclusions and policy implications

With comparatively few ex-ante empirical studies of social indicators in the modelling literature, this paper bridges ex-post and ex-ante modelling work for two selected EU focused social indicators in the domains of education and gender (in)equality. Starting with the ex-post work, a literature review and *a priori* expectations of potential drivers are used to **build** a comprehensive panel database of candidate independent variables. Employing an innovative double-lasso technique **(i.e. a** machine learning **approach for** data selection and reduction techniques), the statistically most pertinent driver elasticities are **selected and** estimated for both indicators.

The econometric results reveal that economic growth, lower agricultural employment, higher education expenditure and trade openness, raise post-secondary education attainment rates amongst 25–29-year-olds. In terms of the absolute magnitude, trade openness is found to have the strongest impact, followed closely by **the share of** education expenditure share**. The** gender wage gap **is narrowed by factors such as** economic growth, higher agricultural employment, a reduced concentration of income held by the richest and poorest 10 percent of the population. **Furthermore,** fewer urban agglomerations and, with the strongest absolute impact, increases in average years of education **emerge as the most powerful driver of** gender pay gap equality.

These findings carry clear policy implications for the European Union. On the one hand, **post-secondary** education is driven by structural factors which directly influence household resources,

ARTICLE IN PRESS

public investment capacity and labour-market demand for high skill workers. The strong positive influence of GDP per capita, education expenditure and trade openness on tertiary attainment reinforce the importance of EU strategies that promote economic resilience, sustained public investment in human capital and deeper integration into global markets. **In the case of trade openness**, it is therefore expected that envisaged trade agreements with Mercosur and India **can boost post-secondary** education **by** increas**ing** high skilled labour demand. Similarly, rises in education expenditure will have positive effects on tertiary education, highlighting the relevance of public spending and commitment to human capital (European Commission's Education and Training Monitor, 2024; 2025). **Conversely,** the negative association between agricultural employment and education attainment highlights the relevance of cohesion policies **designed to** support structural transformation and skill upgrading in rural regions.

On the other hand, the gender wage gap is shaped by a more complex set of institutional, behavioural, labour market dynamics, spatial effects and inequality factors. Policies that raise the level of education have positive effects on the gender wage gap. The strong negative elasticity associated with mean years of schooling further underscores the importance of lifelong learning and upskilling initiatives as tools for narrowing gender disparities. **Beyond education**, the model emphasises that the identified drivers, particularly income distribution, urban agglomeration and employment in agriculture, underscore the need for targeted labour market and social policies. This **supports the integration of** enforceable policies (e.g., flexible working arrangements, conditions of job security and pay, improved public transport, work-life balance directives) that incentivise gender mainstreaming (i.e. consider the systematic incorporation of a gender perspective in all policy processes) into urban planning and infrastructure development (OECD, 2025). Finally, **because income inequality** disproportionately **harms female earnings, the findings emphasize the need for** redistributive and counter-cyclical fiscal policies to **prevent the** gender wage gap **from widening during economic downturns**.

To integrate the ex-post findings into a forward-looking framework, the econometric elasticity estimates are coded into an ex-ante computable general equilibrium (CGE) simulation model. **This model utilizes 2017** starting values for each social indicator and a mixture of published (KC et al., 2024) and in-house projections for both social indicators to 2040 to permit a calibration of their projected baseline trends**.** A global slowdown scenario explores an assumed accumulated region-wide relative reduction of 5% in productivity by 2040 **compared to** the baseline. **Under these conditions, the model predicts that** post-secondary educational attainment rate amongst 25-29 years old falls -4.830%, largely driven by falling per capita GDP, rising agricultural employment rates and reduced education expenditure. Under such macroeconomic uncertainty, policy recommendations could include a commitment to counter-cyclical maintenance of post-secondary education expenditure, targeted support for disadvantaged groups (e.g., means-tested grants, or student loans), and continued labour market

ARTICLE IN PRESS

reforms that promote flexible study arrangements**. Furthermore, g**iven the role of agricultural employment, efforts should be made to strengthen the integration of post-secondary education with rural economies **through** lifelong learning and vocational pathways consistent with the EU's 'Vision' for agriculture (EC, 2025).

The CGE slowdown experiment also reports a relative rise in the EU-wide gender wage gap by 21.577%, which underlines the vulnerability of women in the workforce **during** economic downturns. With relative falls in education playing a key role in this development, targeted support aimed at raising female participation is a priority (e.g., social protection measures for young mothers or underprivileged females in education). The rise in agricultural labour share reduces the gender wage gap, **which** highlights **how** sectoral composition **can** shape wage outcomes. **Consequently**, policy design should promote gender equitable wage setting, career protection and progression in all sectors to ensure that sectoral transitions associated with economic downturns do not **disproportionately** disadvantage female workers.

This research represents an exploratory step toward improving insights on future social indicator trends **while acknowledging certain** assumptions and limitations. A key **limitation** is that behavioural elasticity drivers of social indicators are naturally based on past observations. As such, the econometric response parameters are governed by past conditions and therefore within the ex-ante simulation model, remain structurally invariant to unforeseen future global events. This parametric critique can be commonly levelled at **all** the extraneously estimated behavioural response parameters (i.e., demand and supply elasticities) within a simulation model. This **observation** is particularly **prescient** in the context of rapid technological change or institutional reform, which may alter behavioural responses in ways not captured by historical data. Finally, in our study, the social indicator results do not feedback into the CGE model mechanisms. For example, sustained higher rates of post-secondary attainment would be reflected in the skill ratio of the workforce, which over time would be expected to affect labour productivity, economic performance and structural change.

This research highlights that while the number of candidate drivers for social indicators is expanded through a comprehensive panel database, certain "qualitative" factors remain difficult to measure. These include legal, institutional, cultural, lifestyle factors, and psychological factors such as attitudes towards risk, competition, negotiation (Olivetti and Petrongolo, 2016, Azmat and Petrongolo, 2014), which cannot be easily measured and analysed in an ex-post model, nor implemented within an ex-ante model.[xiii] These omitted dimensions may therefore contribute to unexplained cross-country heterogeneity in both the econometric and simulation results. For this reason, the use of an error term on the CGE model **(as discussed in section 4)** not only compensates for structural differences between the econometric and CGE approaches, but also implicitly internalises these additional considerations.

Despite these shortcomings, the approach offers insights on the known drivers of social indicators and subsequently the potential synergies and trade-offs arising from different uncertainty pathways. Wider coverage in terms of countries, time span and additional variables would be useful for the

ARTICLE IN PRESS

econometric analysis and further extensions of CGE models to accommodate additional drivers reflecting psychological or institutional factors would further benefit research in this area.

ARTICLE IN PRESS

ARTICLE IN PRESS

Ethics declarations

Competing interests

The authors declare no competing interests.

Ethical approval

This article does not contain any studies with human participants performed by any of the authors.

Informed consent

Informed consent was not applicable as the study did not involve human participants, and the raw data is publicly available.

Data availability statement

The study only uses secondary data sources. Datasets and code used in the estimation part of the paper **is** publicly available at Zenodo's repository **(https://doi.org/10.5281/zenodo.18630282).**[xiv] We combine available data sources from Wittgenstein Centre for Demography and Global Human Capital (see Lutz et al., 2018, KC et al. 2024), the World Development Indicators database (World Bank, 2025) and Eurostat SDGs database. Based on the available sources, we determine the sample and timeframe. Since the computable general equilibrium model used (MAGNET) is not open source, we are unable to share this **data**. The key database upon which it is based is the Global Trade Analysis Database (GTAP), which is available at cost from their website.[xv] **Notwithstanding,** the projections of the social indicators data implemented into the CGE model to calibrate our baseline have been included in the repository.

ARTICLE IN PRESS

Acknowledgments

The authors acknowledge support from the BrightSpace project: *Towards a safe and just operating space for EU agriculture. Project GA Nr. 101060075 — HORIZON-CL6-2021-GOVERNANCE-01-12* supported by the European Union.

ARTICLE IN PRESS

References

Acs, S., Ostlaender, N., Listorti, G., Hradec, J., Hardy, M., Smits, P., & Hordijk, L. (2019). Modelling for EU Policy support: impact assessments. *Luxembourg: Publications Office of the European Union*.

Adeleke, O., & McSharry, P. E. (2022). Female enrollment, child mortality and corruption are good predictors of a country's UN education index. *International Journal of Educational Development*, 90, 102561. DOI: 10.1016/j.ijedudev.2022.102561

Aguiar, A., Chepeliev, M., Corong, E., & van der Mensbrugghe, D. (2023). The Global Trade Analysis Project (GTAP) Data Base: Version 11. *Journal of Global Economic Analysis*, *7*(2). https://doi.org/10.21642/JGEA.070201AF (Original work published 2022)

Athey, S., and Imbens, G. W. (2019). Machine learning methods that economists should know about. *Annual Review of Economics*, *11*(1), 685-725. https://doi.org/10.1146/annurev-economics-080217-053433

Autor, D. H., Dorn, D., & Hanson, G. H. (2013). The China syndrome: Local labour market effects of import competition in the United States. *American Economic Review*, 103(6), 2121-2168. DOI: 10.1257/aer.103.6.2121

Azmat, G., & Ferrer, R. (2017). Gender gaps in performance: Evidence from young lawyers. *Journal of Political Economy*, 125(5), 1306-1355. https://doi.org/10.1086/693686

Azmat, G., & Petrongolo, B. (2014). Gender and the labor market: What have we learned from field and lab experiments? *Labour Economics*, *30*, 32-40. https://doi.org/10.1016/j.labeco.2014.06.005

Barro, R. (2013). Education and economic growth. Annals of Economics and Finance, 14(2), 301–328.

Bartelings, H., van den Bos-Verma, M., & van Meijl, H. (2024). Introducing municipal solid waste streams in a CGE framework: A step in the direction of circular economy analysis. *Journal of Global Economic Analysis*, 9(1), 01-44. https://doi.org/10.21642/JGEA.090101AF

Bartoloni, E., Marino, A., Baussola, M., & Romaniello, D. (2023). Urban non-urban agglomeration divide: is there a gap in productivity and wages? *Italian Economic Journal*, 9(2), 789-827. https://doi.org/10.1007/s40797-022-00202-6

Belloni, A., Chernozhukov, V., & Hansen, C. (2014). Inference on Treatment Effects after Selection among High-Dimensional Controls. Review of Economic Studies, 81(2), 608–650. https://doi.org/10.1093/restud/rdt044

Benson, Alan, Danielle Li and Kelly Shue. 2026. ""Potential" and the Gender Promotion Gap." *American Economic Review*, 116 (2): 375–417. DOI: 10.1257/aer.20220831

Benton, T. (2019) Using Scenario Analysis to Address the Future of Food, *European Food Safety Authority (EFSA) Journal*, 17 (51). https://doi.org/10.2903/j.efsa.2019.e170703

Blanchard, E., & Willmann, G. (2016). Trade, education, and the shrinking middle class. *Journal of International Economics*, *99*, 263-278. https://doi.org/10.1016/j.jinteco.2015.10.007

Blau, F. D., & Kahn, L. M. (2017). The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature*, 55(3), 789-865. DOI: 10.1257/jel.20160995

Blau, F. D., Kahn, L. M., Boboshko, N., & Comey, M. (2024). The impact of selection into the labour force on the gender wage gap. *Journal of Labour Economics*, 42(4), 1093-1133. https://doi.org/10.1086/725032

Boulanger, P., Dudu, H., Ferrari, E., & Philippidis, G. (2016). Russian roulette at the trade table: a specific factors CGE analysis of an agri-food import ban. *Journal of Agricultural Economics*, 67(2), 272-291. https://doi.org/10.1111/1477-9552.12156

Boulanger, P. & Philippidis, G. (2015). The EU Budget Battle: Assessing the Trade and Welfare Impacts of CAP Budgetary Reform, *Food Policy*, 51, pp119-130. DOI:10.1016/j.foodpol.2015.01.004

Burfisher, M., 2011. Introduction to Computable General Equilibrium Models, Cambridge University Press. https://doi.org/10.1017/CBO9780511975004

Campagnolo, L., Carraro, C., Eboli, F., Farnia, L., Parrado, R., & Pierfederici, R. (2018). The ex-ante evaluation of achieving sustainable development goals. *Social Indicators Research*, 136, 73-116. https://doi.org/10.1007/s11205-017-1572-x

Campagnolo, L., & Davide, M. (2019). Can the Paris deal boost SDGs achievement? An assessment of climate mitigation co-benefits or side-effects on poverty and inequality. *World Development*, 122, 96-109. https://doi.org/10.1016/j.worlddev.2019.05.015

Caselli, F., & Coleman II, W. J. (2001). The US structural transformation and regional convergence: A reinterpretation. *Journal of Political Economy*, 109(3), 584-616. https://doi.org/10.1086/321015

ARTICLE IN PRESS

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., & Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters. https://doi.org/10.1111/ectj.12097

Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, et al. 2025. "V-Dem [Dataset Version] Dataset v15". Varieties of Democracy (V-Dem) Project. https://www.v-dem.net/data/the-v-dem-dataset. https://doi.org/10.1177/0192512115622046

Corong, E. L., Hertel, T. W., McDougall, R., Tsigas, M. E., & Van Der Mensbrugghe, D. (2017). The standard GTAP model, version 7. *Journal of Global Economic Analysis*, 2(1), 1-119.

Correll, S. J., Benard, S., & Paik, I. (2007). Getting a job: Is there a motherhood penalty? *American Journal of Sociology*, 112(5), 1297-1338. https://doi.org/10.1086/511799

Craigwell, R., Bynoe, D., & Lowe, S. (2012). The effectiveness of government expenditure on education and health care in the Caribbean. *International Journal of Development Issues*, 11(1), 4-18. https://doi.org/10.1108/14468951211213831

Darlington, R. B. (1990). *Regression and linear models*. Mcgraw-Hill College.

Dabla-Norris, E., & Kochhar, K. (2019). Closing the gender gap. *Finance and Development*, *56*(1), 8-11.

Daioglou, V., Stehfest, E., Wicke, B., Faaij, A., & van Vuuren, D.P. (2015). Projections of the availability and cost of residues from agriculture and forestry. *Global Change Biology*, 8, 456-470. https://doi.org/10.1111/gcbb.12285

DG RTD 2017. Research and Innovation perspective of the mid- and long-term potential for advanced biofuels in Europe: Deliverable 1.2 Research and Innovation scenarios for biomass potential, European Commission, Directorate General for Research and Innovation. Brussels.

Diamanti, G., & Duncan, J. (2023). *Gender and agriculture: policy tensions behind the EU gender gap*: summary report of a systematic literature review of the gender responsiveness of Europe's agriculture and rural policy. https://research.wur.nl/en/publications/gender-and-agriculturepolicy-tensions-behind-the-eu-gender-gap-su

Dridi, M. (2014). Corruption and education: Empirical evidence. *International Journal of Economics and Financial Issues*, 4(3), 476-493. https://www.econjournals.com.tr/index.php/ijefi/article/view/781

England, P., Levine, A., & Mishel, E. (2020). Progress toward gender equality in the United States has slowed or stalled. *Proceedings of the National Academy of Sciences*, 117(13), 6990-6997. https://doi.org/10.1073/pnas.1918891117

European Institute of Gender Equality (2017). *Economic Benefits of Gender Equality. 2017. Economic Benefits of Gender Equality in the European Union*: Report on the Empirical Application of the Model: Publications Office of the European Union

European Commission: Joint Research Centre, *Joint Research Centre work programme 2025-2027*, Publications Office of the European Union, 2024, https://data.europa.eu/doi/10.2760/4668727

European Commission (2025): Directorate-General for Education, Youth, Sport and Culture, *Education and training monitor 2025 – Comparative report*, Publications Office of the European Union, 2025, https://data.europa.eu/doi/10.2766/2221794

European Commission (2024) Education and Training Monitor 2024. Luxembourg: Publications Office of the European Union. https://education.ec.europa.eu/education-and-training-monitor.https://op.europa.eu/webpub/eac/education-and-training-monitor/en/comparative-report/comparative-report.html

European Commission (2025). Report on gender equality in the EU. https://euagenda.eu/publications/2025-report-on-gender-equality-in-the-eu

EC (2025) Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, A Vision for Agriculture and Food Shaping together an attractive farming and agri-food sector for future generations COM/2025/75 final

Fernández, R., Isakova, A., Luna, F., & Rambousek, B. (2021). *Gender equality and inclusive growth*. International Monetary Fund.

Gatto, A., Kuiper, M., & van Meijl, H. (2023). Economic, social and environmental spillovers decrease the benefits of a global dietary shift. *Nature Food*, 4(6), 496-507. DOI: 10.1038/s43016-023-00769-y

Gebers, Kathrin et al. (ed.) (2024). *Sustainable Development in the European Union. Monitoring Report on Progress Towards the SDGs in an EU Context*. [eurostat Statistical books 2024 edition]. Luxembourg: Publications Office of the European Union.

ARTICLE IN PRESS

ARTICLE IN PRESS

Gerdes, C. (2011). Using “shares” vs. “log of shares” in fixed-effects estimations. *Journal of Economics and Econometrics*, 54(1), 1-6.

Goldin, C. (2014). A grand gender convergence: Its last chapter. *American Economic Review*, 104(4), 1091-1119. DOI: 10.1257/aer.104.4.1091

Haltiwanger, J., Hyatt, H. R., & Spletzer, J. R. (2024). Rising top, falling bottom: Industries and rising wage inequality. *American Economic Review*, *114*(10), 3250-3283. DOI: 10.1257/aer.20221574

Hasegawa, T., Fujimori, S., Havlik, P., Valin, H., Bodirsky, L.B. et al. (2018). Risk of increased food insecurity under stringent global climate change mitigation policy. *Nature Climate Change*. 8: 699–703. https://doi.org/10.1038/s41558-018-0230-x

Herrendorf, B., & Schoellman, T. (2018). Wages, human capital, and barriers to structural transformation. *American Economic Journal: Macroeconomics*, *10*(2), 1-23. DOI: 10.1257/mac.20160236

Hirukawa, M., Liu, D., Murtazashvili, I., & Prokhorov, A. (2023). DS-HECK: double-lasso estimation of Heckman selection model. *Empirical Economics*, *64*(6), 3167-3195. https://doi.org/10.1007/s00181-023-02406-w

Inglehart, R., C. Haerpfer, A. Moreno, C. Welzel, K. Kizilova, J. Diez-Medrano, M. Lagos, P. Norris, E. Ponarin & B. Puranen et al. (eds.). 2014. World Values Survey: All Rounds - Country-Pooled Datafile Version: https://www.worldvaluessurvey.org/WVSDocumentationWVL.jsp. Madrid: JD Systems Institute.

James, Gareth, Daniela Witten, Trevor Hastie, Robert Tibshirani et al. (2013). *An introduction to statistical learning*, Volume 112: Springer.

Jayachandran, Seema. (2015). The roots of gender inequality in developing countries. *Annual Review of Economics*. 7: 63–88. https://doi.org/10.1146/annurev-economics-080614-115404

KC, S., Dhakad, M., Potančoková, M., Adhikari, S., Yildiz, D., Mamolo, M., Sobotka, T.,Zeman, K., Guy, A., L. Wolfgang, Goujon, A. (2024). Updating the Shared Socioeconomic Pathways (SSPs) Global Population and Human Capital Projections. Working Paper. Laxenburg, Austria: WP-24-003

Keramidas, K., Fosse, F., Diaz-Vazquez, A., Schade, B., Tchung-Ming, S., Weitzel, M., Vandyck, T. & Wojtowicz, K. (2021). Global energy and climate outlook 2020: a new normal beyond Covid-19. Publications Office of the European Union, 10, 608429. DOI: 10.2760/608429 (online)

Kunze, A. (2018). The gender wage gap in developed countries. *The Oxford handbook of women and the economy*, *4*, 369-394.

Leythienne, D., & Pérez-Julián, M. (2021). Gender pay gaps in the European Union. A statistical analysis, Brussels: Eurostat.

Liu, Sitian, and Yichen Su. (2024). The geography of jobs and the gender wage gap. *Review of Economics and Statistics*. 106: 872–881. https://doi.org/10.1162/rest_a_01188

Lutz, Wolfgang, Anne Goujon, Samir Kc, Marcin Stonawski, and Nikolaos Stilianakis. (2018). *Demographic and human capital scenarios for the 21st century*, 2018 assessment for 201 countries: Publications Office of the European Union.

MacKinnon, D. P., Krull, J. L., & Lockwood, C. M. (2000). Equivalence of the mediation, confounding and suppression effect. *Prevention Science*, 1, 173-181.

Mandel, H., & Rotman, A. (2021). Revealing the concealed effect of top earnings on the gender gap in the economic value of higher education in the United States, 1980–2017. *Demography*, *58*(2), 551-570. https://doi.org/10.1215/00703370-9009367

Manning, A., and Petrongolo B. (2008). The part-time pay penalty for women in Britain. *The Economic Journal*. 118: F28–F51. https://doi.org/10.1111/j.1468-0297.2007.02115.x

Moreno, J., Campagnolo, L., Boitier, B. *et al.* (2024) The impacts of decarbonization pathways on Sustainable Development Goals in the European Union. *Commun Earth Environ* 5, 136. https://doi.org/10.1038/s43247-024-01309-7

Moyer, J. D., & Bohl, D. K. (2019). Alternative pathways to human development: Assessing trade-offs and synergies in achieving the Sustainable Development Goals. *Futures*, 105, 199-210. https://doi.org/10.1016/j.futures.2018.10.007

Moyer, J. D., & Hedden, S. (2020). Are we on the right path to achieve the sustainable development goals?. *World Development*, *127*, 104749. https://doi.org/10.1016/j.worlddev.2019.104749

Nisic, N. (2017). Smaller differences in bigger cities? Assessing the regional dimension of the gender wage gap. *European Sociological Review*, 33(2), 292-3044. https://doi.org/10.1093/esr/jcx037

OECD (2025). Gender equality in a changing world. Publishing. https://www.oecd.org/gender/

ARTICLE IN PRESS

Olivetti, C., & Petrongolo, B. (2016). The evolution of gender gaps in industrialized countries. Annual Review of Economics, 8(1), 405-434. https://doi.org/10.1146/annurev-economics-080614-115329

O'Neill, B. C., Kriegler, E., Ebi, K. L.; Kemp-Benedict, E; Riahi, K. et al (2017). The roads ahead: Narratives for shared socioeconomic pathways describing world futures in the 21st century. Global Environmental Change. 42: 169–180. https://doi.org/10.1016/j.gloenvcha.2015.01.004

Orbons, K.; van Vuuren, D.P., Ambrosio, G., Kulkarni, S., Weber, E., Zapata, V., Daioglou, V., Hof, A.F., Zimm, C. (2024) A review of existing model-based scenarios achieving SDGs: progress and challenges, Global Sustainability, 7, 1-7. DOI: https://doi.org/10.1017/sus.2023.20

Philippidis, G., Álvarez, R.X., Di Lucia, L, Hermoso, H. G., Martinez, A. G., et al. (2024). The development of bio-based industry in the European Union: A prospective integrated modelling assessment. *Ecological Economics*. 219: 108156. https://doi.org/10.1016/j.ecolecon.2024.108156

Philippidis, G., Shutes, L., M'Barek, R., Ronzon, T., Tabeau, A., & Van Meijl, H. (2020). Snakes and ladders: World development pathways' synergies and trade-offs through the lens of the Sustainable Development Goals, *Journal of Cleaner Production*. 267: 122147. https://doi.org/10.1016/j.jclepro.2020.122147

Piketty, T., Saez, E., & Zucman, G. (2018). "Distributional national accounts: methods and estimates for the United States, *The Quarterly Journal of Economics*. 133: 553–609. https://doi.org/10.1093/qje/qjx043

Quadlin, N., Van Heuvelen, T., & Ahearn, C.E. (2023). Higher education and high-wage gender inequality, *Social Science Research*. 112: 102873. https://doi.org/10.1016/j.ssresearch.2023.102873

Rajkumar, A.S., & Swaroop, V. (2008). Public spending and outcomes: Does governance matter? *Journal of Development Economics*. 86: 96–111. https://doi.org/10.1016/j.jdeveco.2007.08.003

Riahi, Keywan, Detlef P Van Vuuren, Elmar Kriegler et al. (2017). The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview, Global *Environmental Change*. 42: 153–168. https://doi.org/10.1016/j.gloenvcha.2016.05.009

Rockström, J., Steffen, W., Noone, K., Persson, Å., Chapin III, F. S., Lambin, E., et al (2009). Planetary boundaries: exploring the safe operating space for humanity. *Ecology and society*, *14*(2). https://www.jstor.org/stable/26268316

Schuh, B, Maucorps, A., Münch, A., Brkanovic, S., Dwyer, J. et al. (2019). *Research for AGRI Committee – The EU farming employment: current challenges and future prospects*, European Parliament, Policy Department for Structural and Cohesion Policies, Brussels

Senadza, B., & Hodey, L.S., (2015). Effect of public education expenditure on selected primary education outcomes in Sub-Saharan Africa. *International Journal of Education Economics and Development*. 6: 314–330. https://doi.org/10.1504/IJEED.2015.075794

UNESCO GEM Report (2025). Global Education Monitoring Report 2025: Gender report. Paris: UNESCO. https://www.unesco.org/gem-report

United Nations. 2000. Resolution adopted by the General Assembly. United Nations Millennium Declaration (A/55/L.2). https://docs.un.org/en/A/RES/55/2

United nations. 2015. The Millennium Development Goals Report. Retrieved from https://www.un.org/millenniumgoals/2015_MDG_Report/pdf/MDG%202015%20rev%20(July%201).pdf

Urhie, E., Ewetan, O.O., & Okodua, H. (2014). National Income, Public Expenditure on Education and Educational Attainment: A Comparative Analysis. *Covenant Journal of Business and Social Sciences* 6(2), 1-13. https://journals.covenantuniversity.edu.ng/index.php/cjbss/article/view/12

Urminsky, O., Hansen, C., & Chernozhukov, V. (2016). Using double-lasso regression for principled variable selection. Available at http://dx.doi.org/10.2139/ssrn.2733374.

Van Meijl, H., Shutes, L., Valin, H., Stehfest, E., van Dijk, M., Kuiper, M., et al (2020). Modelling alternative futures of global food security: Insights from FOODSECURE. *Global Food Security*, 25, 100358. https://doi.org/10.1016/j.gfs.2020.100358

Woltjer, G. B., Kuiper, M., Kavallari, A., van Meijl, H., Powell, J. et al (2014). The MAGNET model. Module description (No. 14-57). LEI. https://edepot.wur.nl/310764

World Bank (2025) World Development Indicators, https://databank.worldbank.org/source/world-development-indicators

ARTICLE IN PRESS

Endnotes

[i] For example, after childbirth women decide whether they return to work, whether it **be** full-time or part-time. **This decision can introduce** self-selection bias as women that return to work may have different characteristics that are not observed by the researcher, such as higher motivation, better skills or a stronger support network. Similarly, there can be simultaneity bias as the decision to return to **work,** and the labour market outcomes (e.g. wages, career progression) can influence each other simultaneously.

[ii] Regarding performance, measured using detailed individual data on billed hours and client revenue, Azmat and Ferrer (2017) find that in the legal profession, part of the gender gap in earnings is explained by women billing fewer hours and acquiring less client revenue than men.

[iii] See online resource A.1 **including Table A1 for more details.**

[iv] The coefficient of interest may be either artificially strong (when the covariate is a confound) or may be artificially weak (a suppression effect; MacKinnon et al. 2000).

[v] Endogeneity coming from measurement error or unobserved factors **is** still **as** relevant as in any ordinary least squares regression setting.

[vi] There are some slight differences as they come from different databases, thus we introduce a constant from 2020 to 2040 that collects the difference between the value in Eurostat and the SSPs projections for GDP in 2020.

[vii] Only for projection purposes, lasso requires information for all the controls in the model, **any missing information for specific countries is replaced to ensure a complete set of controls.** See Table A4 in Appendix A2 for more details. Although this model predicts accurately on average for each country, **predicted values** may **differ with** the last observed **year (**2020**), which is** the starting point for the linear prediction. **Therefore**, we adjust the country-specific projections by the observed error in 2020 (i.e. gap between observed and fitted value). Similarly, the projections of **the** gender wage gap, are also adjusted by the error observed in 2020.

[viii] A fuller description of the CGE model approach is provided in online resource C. A further description of **the** MAGNET **model** is also available online at https://www.magnet-model.eu/.

[ix] Calibration ensures that for a given equilibrium set of starting prices, the model equations reproduce perfectly and consistently the underlying equilibrium transaction data structures for final, intermediate and investment demand in the starting period data.

[x] **The online resource C provides further details including the regional and commodity disaggregation (Table C1) and the description of the baseline driver shocks employed in the study (Table C2).**

[xi] In the case of the gender wage gap indicator, the inserted equation excludes the (statistically significant) driver of urban agglomeration, as no such driver (currently) exists in MAGNET.

[xii] Trade openness is measured as the ratio of the change in the value of a region's trade relative to its GDP compared with the change in the value of global trade compared with its GDP.

[xiii] Thanks to an anonymous reviewer, **the study explores** the inclusion of additional variables regarding **the** women**'s** political empowerment index or variables capturing the egalitarian component of democracy from **the** VDEM database (Coppedge et al. 2025). Similarly, we also explore the inclusion of Survival vs. Self-Expression values from the Inglehart–Welzel Cultural Map (IWCM) database (Inglehart et al, 2014) from the World values Survey *https://www.worldvaluessurvey.org.* However, **these variables were ultimately excluded from the results** due to the methodological constrains of **the** lasso **technique** which cannot select from re**gressors** generated **elsewhere** via Bayesian analysis or second order latent variables.

[xiv] **Database citation: Garcés, I., Philippidis, G., Sanjuán, A. I., & Xavier Álvarez, R. (2026). Bridging ex-post and ex-ante EU social indicator trends: An explorative modelling analysis for SDG4 and SDG5 [Data set]. BrightSpace Horizon Europe project GA Nr. 101060075.**

[xv] Please see additional information and access the GTAP database at *https://www.gtap.agecon.purdue.edu/databases/default.asp*