European Commission

JRC TECHNICAL REPORT

# An update on the forecast of Europe 2020 headline targets on education and training

Esperanza Vera-Toscano

2015

EUR 26077 EN

Joint Research Centre

European Commission
Joint Research Centre
Deputy Director-General Office, Econometrics and Applied Statistics Unit

Contact information
Esperanza Vera-Toscano
Address: Joint Research Centre, Via Enrico Fermi 2749, TP 361, 21027 Ispra (VA), Italy
E-mail: esperanza.vera-toscano@jrc.ec.europa.eu
Tel.: +39 0332 78 5103

JRC Science Hub
https://ec.europa.eu/jrc



JRC97601

EUR 26077 EN

ISBN 978-92-79-51831-7

ISSN 1831-9424

doi:10.2788/05658

Luxembourg: Publications Office of the European Union, 2015

## 1. Introduction[1]

In Europe, educational attainment has steadily increased since the beginning of the 20th century and, likewise, the share of highly-educated individuals over the past few decades. However, while Europe overall is likely to achieve its targets on education and training within the Europe 2020 strategy, there is growing divide between best and low performing countries that raises doubts with respect to economic convergence prospects. This technical brief builds on the work undertaken by Dragomirescu-Gaina and Weber (2013), aiming to provide an update to the question whether Europe as a whole and each of its Member States will reach the twofold Europe 2020 target on early leavers from education and training (ELET) and tertiary education attainment (TEA) by 2020[2].

The models are built on the theoretical framework of human capital using a panel structure. Forecasts are constructed under simple but realistic assumptions about the expected level of adults' educational attainment, given the determinants of schooling decisions uncovered by our empirical analysis. Our results confirm those provided by Dragomirescu-Gaina and Weber (2013). As stated by them, these forecasts tell us how ELET and TEA are likely to develop over the next years if nothing changes in terms of policy measures which provides interesting information for policy action, especially for those countries where the expected developments of model's determinants are not enough to foresee a positive outcome. This technical brief is organised as follows. Section 2 provides some background information on the model specifications chosen for each of the Europe 2020 targets briefly highlighting the implications of using a given specification with updated data downloaded in different time periods. Section 3 provides information about the data used and the econometric results of the regression models both using the time frame of the previous forecast (up to 2012) and the updated one (up to 2014). Forecast results are shown in Section 4. Finally, Section 5 concludes.

## 2. Background information on the model specifications chosen

There is a vast socio-economic literature on the economics of human capital investment. At large, it is said that individuals' expectations on future earnings, based on the information available when the decision is made, are important determinants in deciding on the amount of education with the highest expected return. However, since a measure of subjective expectations is not always available, we need to rely on empirical proxies of them. As stated by Dragomirescu-Gaina and Weber (2103) and Dragomirescu-Gaina, Weber and Elia (2015), it is commonly assumed that a good approximation of the expected return to education is the difference in earnings from undertaking and not undertaking education (see Acemoglu and Pischke, 2001). Autor et al. (1998) and Acemoglu (2002) among many others suggest the existence of a positive feedback loop between labour income and skills on the back of productivity advancements. Their evidence shows that technology shifts over recent decades have

[1] This technical work benefited from discussions and valuable comments from Paolo Paruolo and Leandro Elia. Their contributions are very much appreciated.

[2] The two Europe 2020 headline targets are also benchmarks in the strategic framework for European cooperation in education and training (ET 2020). See COM(2015) 408 final and SWD(2015) 161 final.

favoured skilled or highly educated workers. For the particular case of ELET, which usually lack skills and face poor employment prospects, their planning horizon is more limited in time and their discount rate is higher. Therefore, getting a (first) job would be more important to them than the longer term labour income stream. Thus, Pissarides (1981) was among the first to observe the cyclical component of dropout rates and the myopic reaction to cyclical swings in economic activity. In a more recent study using cross sectional data, Petrongolo and Segundo (2002) found clear evidence of youth unemployment driving staying-on rates in Spain, after accounting for family background. Thus, not only differences in earnings but employment prospects are considered optimal determinants of individual's education investment.

A second major theoretical strand dates back to Becker and Tomes (1979, 1986) who consider the role of family socio-economic background in the education investment process. Utility is maximized across generations, with family acting as the central decision maker in this process. Both theoretical and empirical literature reveals a positive relationship between family income and children's education attainment (Cameron and Heckman 1998; Acemoglu and Pischke, 2001; Black and Devereux, 2011).

Consistent with the above-mentioned literature, Dragomirescu-Gaina and Weber (2013) considered the specification for ***early leavers from education and training*** as a function of parents' education together with employment prospects. They used parents' education instead of income. Unlike earnings, education has some advantages in terms of estimating intergenerational transmission (see Black and Devereux, 2011). First, measurement issues are less of a worry given that, once completed, education is not subject to transitory shocks or life-cycle movements. Second, parents' education might better reflect family permanent income. Indeed, Cameron and Heckman (1998) find that permanent instead of current family income has the key role in explaining the impact of financial constraints on children education attainment. Further, they expose the link between education decisions and employment opportunities by using unemployment rate as a summary of the available set of opportunities outside the education system.[3]

Likewise, their specification for ***tertiary education attainment*** includes parent's education attainment as a proxy for family background and labour productivity as a proxy for expected labour payoff.

These were the simple but robust model specifications chosen by the authors so as to provide long-term forecasts for the Europe 2020 targets, nonetheless, it is worth highlighting that they were selected after a thorough battery of model selection and specification tests. First, the appropriate lag structure of the determinants of TEA and ELET was chosen by using Akaike (AIC) and Bayesian (BIC) information criteria; second, test were undertaken to control for lack of residual autocorrelation using a number of tests proposed by the literature (Arellano-Bond, 1991; Baltagi-Wu, 1999; Wooldridge, 2002); third,

[3] Dragomirescu-Gaina and Weber (2013) considered different specifications based on the theoretical framework for human capital investment. Further, they considered a battery of model selection and specification tests to reach simple but robust model specifications for each benchmark. Thus, these econometric models are just modest representations of the reality, country specific judgements can be added to improve the message of the forecast and therefore alter our qualitative evaluation. As an example, the UK and Luxembourg have big gaps in the education attainment for two consecutive cohorts, probably due to high migration. While migration flows were not included in this study, alternative scenarios could nevertheless consider different projections for parent' education rather than those included here, depending on the assumptions behind migration flows.

coefficients stability was also tested by varying the estimation time-sample and; fourth, forecasting accuracy was evaluated using out-of-sample root mean square errors (RMSE) 1 to 4 years ahead.

Another possibility would have been that, since progress vis-à-vis these Europe 2020 targets is increasing/decreasing gradually over time (see Figures 1 and 2 later in the document), it might have been appropriate to fit a sloping line. That is, a **linear trend model** which is a special case of a simple regression model in which the independent variable is just a time index variable, i.e., 1, 2, 3, etc. (see country specific charts in Appendix 1 comparing our preferred regression model and the linear trend model). While the data may argue in favor of the linear trend model for some countries (e.g. Greece, Italy or Ireland for ELET and Denmark or Cyprus for TEA), empirical research shows that they are often inappropriate for socio-economic data. Most naturally occurring socio-economic time series do not behave as though there are straight lines fixed in space that they are trying to follow: real trends change their slopes and/or their intercepts over time. The linear trend model tries to find the slope and intercept that give the best average fit to all the past data, and unfortunately its deviation from the data is often greatest near the end of the time series, where the forecasting value is. Moreover, a multiple regression model may be more appropriate when trying to draw some policy recommendations on the achievement of these headline targets on education and training over time.

### 2.1. Model specification for tertiary education attainment

The authors consider ***tertiary education attainment*** for a reference population cohort (aged 30-34), denoted by $\mathrm{H}$, as a function of both family socio-economic background, $\mathrm{F}$, and expected labour market payoff, $\mathrm{E}$. All three variables $\mathrm{H}, \mathrm{F}$ and $\mathrm{E}$ can carry two indexes: an age and a time index. To see how the education level of individuals in a given age bracket evolves over time, one needs to fix the age index, $\mathrm{g}$, and let the time index, $\mathrm{t}$, vary. Assuming a constant distance between child's and parents' age, we can characterize both $\mathrm{H}$ and $\mathrm{F}$ using the same age index, $\mathrm{g}$. Yet, the observed education attainment $\mathrm{H}$ is the result of a decision process that occurred in the past, at an age $\mathrm{g}^0 < \mathrm{g}$, the age when the education attainment is measured. As such, the information set available when formulating expectations for future income should correspond to the decision relevant age, i.e. $\mathrm{g}^0$. The following equation represents the basis of our empirical strategy:

$$\mathrm{H}(\mathrm{g})_{\mathrm{c,t}} = \alpha_{\mathrm{c}} + \beta_{\mathrm{c}} * \mathrm{t} + \gamma * \mathrm{F}(\mathrm{g})_{\mathrm{c,t}} + \lambda * \mathrm{E}(\mathrm{g}^0)_{\mathrm{c,t}} + \varepsilon_{\mathrm{c,t}} \qquad (1)$$

Where, $\mathrm{H}(\mathrm{g})_{\mathrm{c,t}}$ represents the population share of young individuals in the age bracket $\mathrm{g}$ having completed tertiary education, as measured in country $\mathrm{c}$ and at time $\mathrm{t}$. The term $\mathrm{F}(\mathrm{g})_{\mathrm{c,t}}$ represents the proxy for family socio-economic background and $\mathrm{E}(\mathrm{g}^0)_{\mathrm{c,t}}$ is the expected labour market payoff. The coefficient $\alpha_{\mathrm{c}}$ is a country-specific constant term meant to capture stable institutional factors affecting education attainment over time while $\beta_{\mathrm{c}}$ is the country-specific time trend, summarizing any consistent and gradual institutional improvements. $\varepsilon_{\mathrm{c,t}}$ is a country- and time-specific error term.

It is worth noting that, $\mathrm{g}^0$ being relevant for the decision-making process remains unobservable. We assume that education choices are based only on the set of information available at the decision-making

age, $g^0$. Within our time-series approach, the unobservable nature of $g^0$ can be overcome by using lags that can empirically approximate the time gap between education decision and education measurement. Dropping g from equation (1), we get:

$$H'_{c,t} = \alpha_c + \beta_c * t + \gamma * F'_{c,t} + \lambda * E'_{c,t-s} + \varepsilon'_{c,t} \qquad (2)$$

where $H'_{c,t}$ is the TEA target, $F'_{c,t}$ is the proxy for family socio-economic background and $E'_{c,t-s}$ is the labour market payoff relevant for individuals aged 30-34. All variables are measured at time t, except $E'_{c,t-s}$ which is the s time-lag of the proxy for expected return (where $s = g - g^0$).

When trying to estimate (2) directly, one needs to address the potential time dependence of the endogenous variable $H'_{c,t}$. This is necessary since the TEA target spans over 5 consecutive cohorts, covering all individuals aged 30 to 34. In this context, an alternative would have been to estimate a dynamic panel, i.e. using difference or system GMM as proposed by Arellano-Bond (1991), Blundell and Bond (1998). Unfortunately, we cannot adopt such estimation strategy given the nature of the benchmark variable and the short time-series available. In particular, given the 5 consecutive cohorts included in the benchmark, one would need to use at least a 5-year lag as instrument to remove correlation between (differenced) error term and the (differenced) lagged dependent variable. This would severely limit sample size and thus weaken the estimation with adverse consequences on the forecasting exercise. In addition, equation (2) cannot be rigorously estimated if the data are non-stationary. Unfortunately, univariate and multivariate unit root tests cannot be applied in our context because of the short time-series, the presence of structural breaks due to methodological changes in collecting the data and the limited number of countries.[4] To tackle the potential non-stationarity in the data, we prefer to specify the model in first-differences:[5]

$$\Delta H'_{c,t} = a + \beta_c + \gamma' * \Delta F'_{c,t} + \lambda' * \Delta E'_{c,t-s} + e_{c,t} \qquad (3)$$

where $\beta_c$ is now interpreted as a country-specific constant term, a is a regression constant and $\Delta$ is the first-difference operator. Equation (3) sets out the dynamics of TEA as a combination of consumption and investment motivations. The lag specification of the labour productivity was not imposed *a-priori* but selected empirically and equal to 13 years lag.

The model was estimated on a balanced sample of 12 EU countries, namely Belgium, Denmark, Finland, France, Germany, Ireland, Italy, the Netherlands, Portugal, Spain, Sweden and the U.K. Together, these countries represented 76% of total EU27 population and generate 88% of its GDP (as of 2012).

[4] Cross sectional dimension N equals at most 28 and time dimension T is 14 annual observations for TEA and 22 for ELET. Asymptotic properties of panel unit root tests require N/T→ 0, which is not met by our dataset. The Im-Pesaran-Shin-type test would be appropriate for fixed N and fixed T but does not allow for serial correlation, which is a major concern in this context, as we will see later. Moreover, multivariate unit root tests require the assumption of independence of the units, which cannot be held in our case.

[5] First-differentiating implies that some of the information contained in the original data would be lost if models' variables were cointegrated. However, this assumption is hardly testable in our data. See Asteriou and Agiomirgianakis (2001) for a model using cointegration techniques on education attainment data.

### 2.2. Model specification for early leavers from education and training

As with the model for TEA, the preferred econometric specification for ***early leavers from education and training***, denoted by $S$, uses total unemployment rate (lagged 6 years) as a proxy for employment prospects $L$, capturing labour market conditions and business cycle dynamics, further including adults' education as a proxy for family background and borrowing constraints $R$. The basis of the empirical strategy is given in the equation below:

$$S'_{c,t} = \alpha_c + \beta_c * t + \gamma * R'_{c,t} + \lambda * L'_{c,t-s} + \varepsilon'_{c,t} \qquad (4)$$

More specifically, the specification is also estimated in first differences with the following general form:

$$\Delta S'_{c,t} = a + \beta_c + \gamma' * \Delta R'_{c,t} + \lambda' * \Delta L'_{c,t-s} + e_{c,t} \qquad (5)$$

The model was estimated on an unbalanced panel for all EU countries.

### 2.3. Update on the information used

Before providing any results on the updating exercise, it is worth keeping in mind some important features of it. First, for both benchmarks, we are using the exact same specifications as those reported by Dragomirescu-Gaina and Weber (2013) and Dragomirescu-Gaina, Weber and Elia (2015), thus we are not questioning at any moment the validity of the models, on the contrary, we consider them as simple but stable and robust models to provide long-term forecasts for the ET benchmarks. However different data download dates[6] are likely to cause changes in the regression estimates. Data is normally updated by data providers, among other reasons, to change provisional values to final ones, to fill in missing values or to correct identified mistakes, this can only result in better estimations of both models. Second, we will run the models both until 2012 and then until 2014 to compare previous and current forecast. The inclusion of two more years in the models plus the use of more accurate data to begin with, as well as the information for the 28th country: Croatia, hopefully will result in the provision of better estimates for these ET2020 targets in 2020.

## 3. Data and Econometric Results

All the historical data used in this technical brief has been taken from the European Statistical Office (Eurostat) or AMECO database of the European Commission maintained by DG ECFIN. Data for the twofold Europe 2020 headline target in education and training is available at country level with an annual frequency being compiled from the EU Labour Force Survey. From an empirical perspective, the data set is limited in the sense that the available time series are short and there are many missing values and/or breaks due to methodological change[7], which represents a drawback for any empirical analysis.

[6] For Dragomirescu-Gaina and Weber (2013) and Dragomirescu-Gaina, Weber and Elia (2015) data was downloaded up to April 2013. For this technical brief data has been downloaded in June 2015.

[7] Data flagged by Eurostat with a (b)

Data for ELET is collected from 1992 to 2014 unless data is not available for a given year and from 2000 to 2014 for TEA (see Figures 1 and 2 below) to have a grasp of data availability and trends per MS and for the EU.

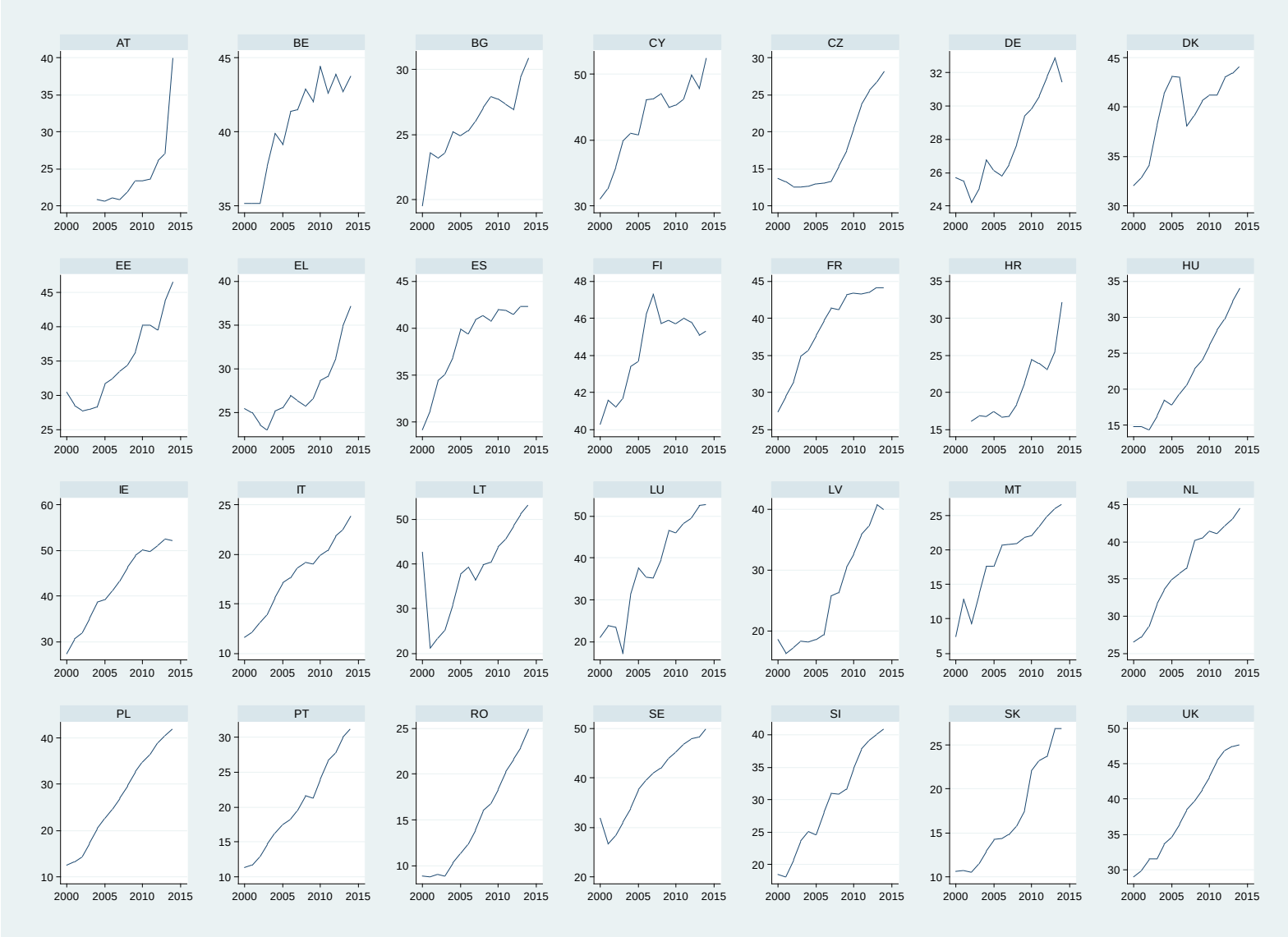
Data on the education attainment of the adult population (vectors F and R in equations 3 and 5), which is used as a proxy for parent's education is also available up to 2014 from Eurostat. More specifically we will use:

- **TEA** specification: share of 55-64 years-old with tertiary education attainment;
- **ELET** specification: share of females 35-44 years-old with less than primary, primary and lower secondary education (levels 0-2) and share of males 45-54 years-old with less than primary, primary and lower secondary education (levels 0-2).

In particular, for ELET, they used adult population education attainment, split by gender, and then they selected the age groups in order to match as closely as possible a typical parental relation between adult cohorts and youth cohorts. The structure of the Labour Force Survey (LFS) does not allow us to exactly associate "parents" with "children" or young individuals included in the EU2020 headline target simply because most of them do not belong to the same household. According to Eurostat, the mean age of women at childbirth was 30 years as of 2011, with a minimum of 27.1 for Bulgaria and a maximum of 31.5 for Ireland and Spain. Statistics further shows that fathers tend to be older than mothers. In simple terms, they followed over the years the educational attainment of children belonging to an average family constructed using the age brackets outlined above.

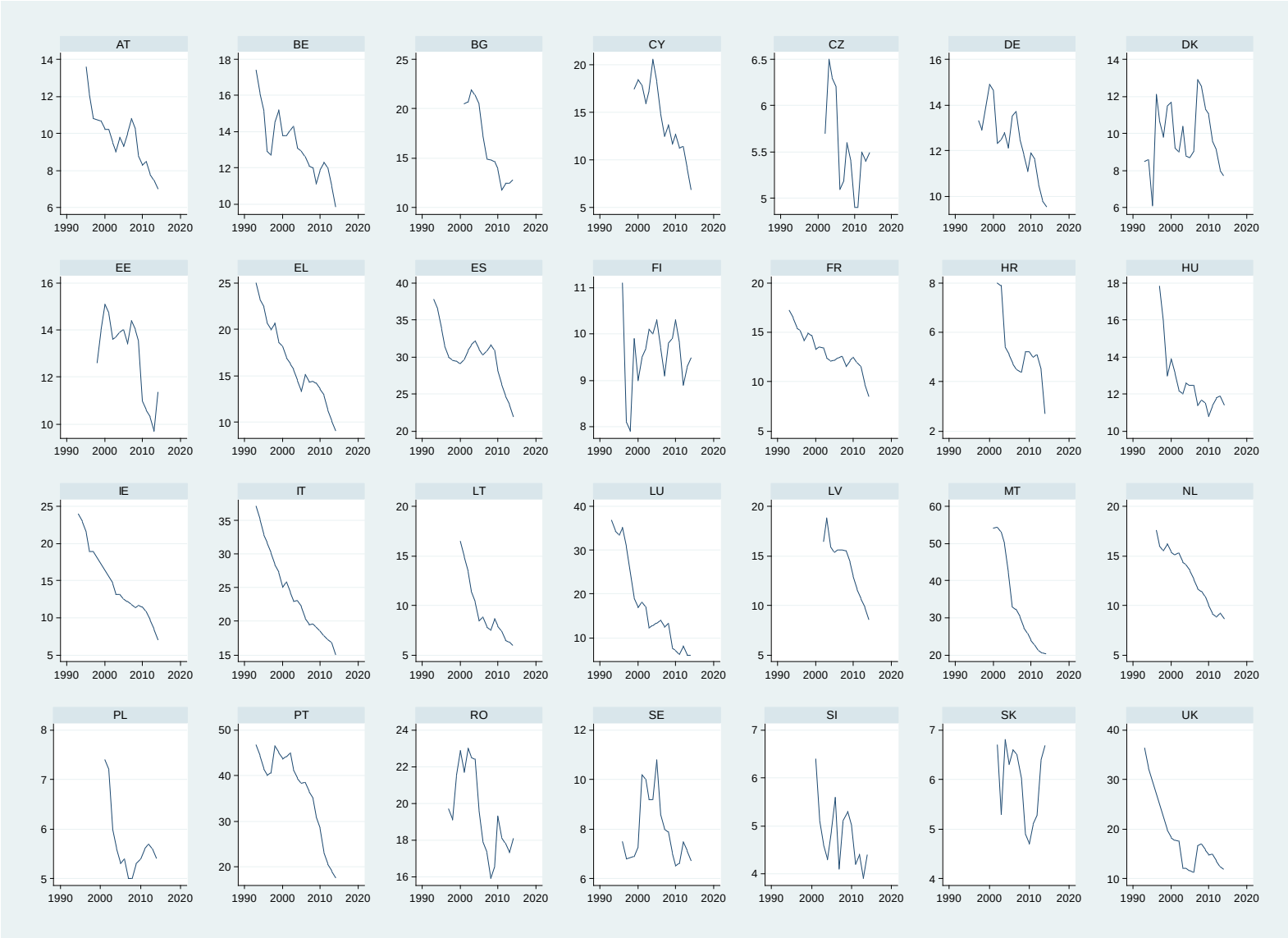
Lastly, data on the economic indicators needed are drawn from Eurostat (annual unemployment rate up to 2014) and AMECO (average annual hours worked per person employed as a measure of labour productivity up to 2014). Given the time lag for these indicators (6 years for unemployment and 13 years for labour productivity) we have to go back as much as possible pending on data availability. The database used is available upon request from the author.

**Figure 1.** TEA by country




Note: Scale of the vertical axis might be different depending on the country. The ET2020 benchmark is set to reach at least 40% of TEA.

**Figure 2.** Share of ELET by country




Note: Scale of the vertical axis might be different depending on the country. The Europe 2020 headline target on ELET is set below 10%

### 3.1. Econometric Results

#### 3.1.1. Results for tertiary education attainment

Table 1 reports the model's estimated coefficients along the specification given by equation (3) for the timespans 2000-2012 (column 1) and 2000-2014 (column 2). Beyond controlling for the share of adults with high education and labour productivity, the model also controls for the major breaks in TEA time-series that occurred around 2003 in most EU countries.[8]

The results of the model are economically and statistically significant and don't change much due to the increase of the sample in 2 years more per country. In particular, the coefficient of 0.43 for the high education adults' share indicates that the share of tertiary educated young people might grow by 0.43 percentage points as a results of a 1% increase in the share of adults with a university degree or higher. These results are consistent with the wealth of empirical investigations that typically find an intergenerational education correlation of the order of 0.3-0.5 for Western Europe (see for instance, Hertz et al., 2007, Chevalier et al., 2003, 2013).

**Table 1**. Estimates of the Δ log share of tertiary educated individuals aged 30-34.

| Model (estimation period) | 2001-2012 | 2001-2014 |
|---|---|---|
| Δ log(adults' share high education, 55-64) t | .43*** | .43*** |
| | (0.06) | (0.06) |
| Δ log(labour productivity), t-13 | 0.48** | 0.47** |
| | (0.24) | (0.22) |
| Constant | 0.028*** | 0.028*** |
| | (0.007) | (0.007) |
| Observations | 144 | 168 |
| *Adj.* $R^2$ | 0.238 | 0.244 |
| # of countries | 12 | 12 |
| Control for breaks in series | yes | yes |
| Year dummies | - | - |
| Country dummies | - | - |

Note. OLS estimates. standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Countries included in the sample are: BE, DE, DK, FI, FR, IT, SE, NL, ES, IE, PT, UK.

Higher labour productivity, our proxy for expected returns, leads to more schooling. A 1% rise in productivity causes an increase in TEA by an amount of 0.48%. The 13$^{th}$ lag of labour productivity when subtracted from the age of the reference population, i.e. 30-34, hints at an interesting overlap with the decision time on enrolment (and/or graduation). The long lag should allay any concerns about potential reverse causality between productivity and higher education attainment in the model.

Findings seem to suggest that higher education attainment is more responsive to expected income than to intergeneration mobility. With rising inequalities over the past decades in many countries, it is not

[8] Dummies were included to control for the presence of breaks in time-series as follows: from 2000 onwards for BE, FI and IT; 2001 for SE; 2002 for PT; 2003 for FR and DK; 2004 for UK; 2005 for DE and ES; 2007 for IE and 2010 for NL.

surprising to see that the decision to invest in higher education today is driven to a larger extent by income motivations (Winchester and Greenaway, 2007).

### 3.1.2. Results for early leavers from education and training

Results for ELET are reported in Table 2 below. These results are pretty much in line with those provided by Dragomirescu-Gaina and Weber (2013). In particular, the coefficients for the share of females' with low education of 0.12 in column 1 indicates that the share of ELET can grow by 0.12 percentage points as a result of a 1% increase in the share of low educated females aged 35-44. Likewise, the coefficient of 0.39 for the low education male adults' share in column 1 indicates that the share of ELET might grow by 0.39 percentage points as a result of a 1% increase in the share of male adults with low education, suggesting in this case that ELET is more responsive to intergeneration mobility. On the contrary, a 1% increase in the unemployment rate is likely to decrease the share of ELET by 0.07 percentage points. Thus, adverse economic conditions prevent people from dropping school. In summary, this model suggests that ELET seems more responsive to family background that to expected labour payoffs.

**Table 2**. Estimates of the Δ log share of ELET aged 18-24.

| Model (estimation period) | 1993-2012 | 1993-2014 |
|---|---|---|
| Δ log(share of females' with low education, age 35-44)$_t$ | 0.12* | 0.20*** |
| | (0.06) | (0.06) |
| Δ log(share of males' with low education, age 45-54)$_t$ | 0.39*** | 0.28*** |
| | (0.07) | (0.07) |
| Δ log(total unemployment rate)$_{t-6}$ | -0.07* | -0.051 |
| | (0.04) | (0.038) |
| Constant | -0.011* | -0.018*** |
| | (0.006) | (0.06) |
| Observations | 346 | 402 |
| *Adj.* $R^2$ | 0.144 | 0.402 |
| # of countries | 28 | 28 |
| Year dummies | - | - |
| Country dummies | - | - |

Note. OLS estimates. standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All 28 EU countries have been included in this regression. Only year dummies before 2000 were included to counter the unbalanced panel specification.

## 4. Forecast results

To provide an answer to our initial question whether the EU27 as a whole will be able to reach the twofold target on education and training by 2020, we need two additional assumptions. First, we assume that the EU28-aggregate behaves according to the best model specification discussed in section 3; with no country-specific terms included in any of the regressions, this assumption should be straightforward. Second, for the particular case of TEA, we argue that the 12 countries in the best model specification provide a good approximation of the EU28-aggregate mainly because they represent a major share in terms of both population and GDP; including more than 12 countries was not supported by their empirical strategy as discussed in Dragomirescu-Gaina, Weber and Elia (2015).

EU28-aggregate and country forecasts are built in two steps. In the first step, we need EU28 and MS projections for the different determinants used in both regressions, i.e. adult education attainment, labour productivity and unemployment rate. For the first set of determinants, that is, share of females with low education (age 35-44), share of males with low education, age 45-54, and share of individuals with higher education, age 55-64 we adopt a simple extrapolation method based on replicating an aging process to obtain projections over the period 2013-2020 (see Appendix A in Dragomirescu-Gaina, Weber and Elia, 2015)[9]. For the labour productivity and unemployment determinants, the long lag used in the estimation allows us to generate forecasts up to 2020 using readily available data. In the second step, we use the model coefficients to compute the expected change in the share of tertiary educated 30-34 year-olds and ELET aged 18-24 up to 2020.[10]

Furthermore, since the forecasts produced under the current approach are **conditional forecasts** we must refresh here the assumptions under which they are valid. They are:

- A "no policy change" scenario, meaning that we are not taking into account any reform that might affect the headline targets over the forecasting horizon, except those that have already produced effects observable in the data used in this study, i.e. data up to 2014. This very much excludes government discretion as a source of uncertainty but also has an immediate policy implication: it suggests that any future policy action that would explicitly target education attainment could still make a difference in some countries.
- The projections of the model determinants are all accurate, so that there is no uncertainty stemming from them. This assumption could allow us to work under different alternative scenarios for parental education, other than the one illustrated in the Annex of Dragomirescu-Gaina et al. (2015).
- The uncertainty reflected in these forecasts is a by-product of the econometric modelling approach we have taken. An econometric model is only a stylized and simplified representation of reality, so there would always be some other determinants or some other transmission channels not accounted for.

### 4.1. TEA and ELET forecast results

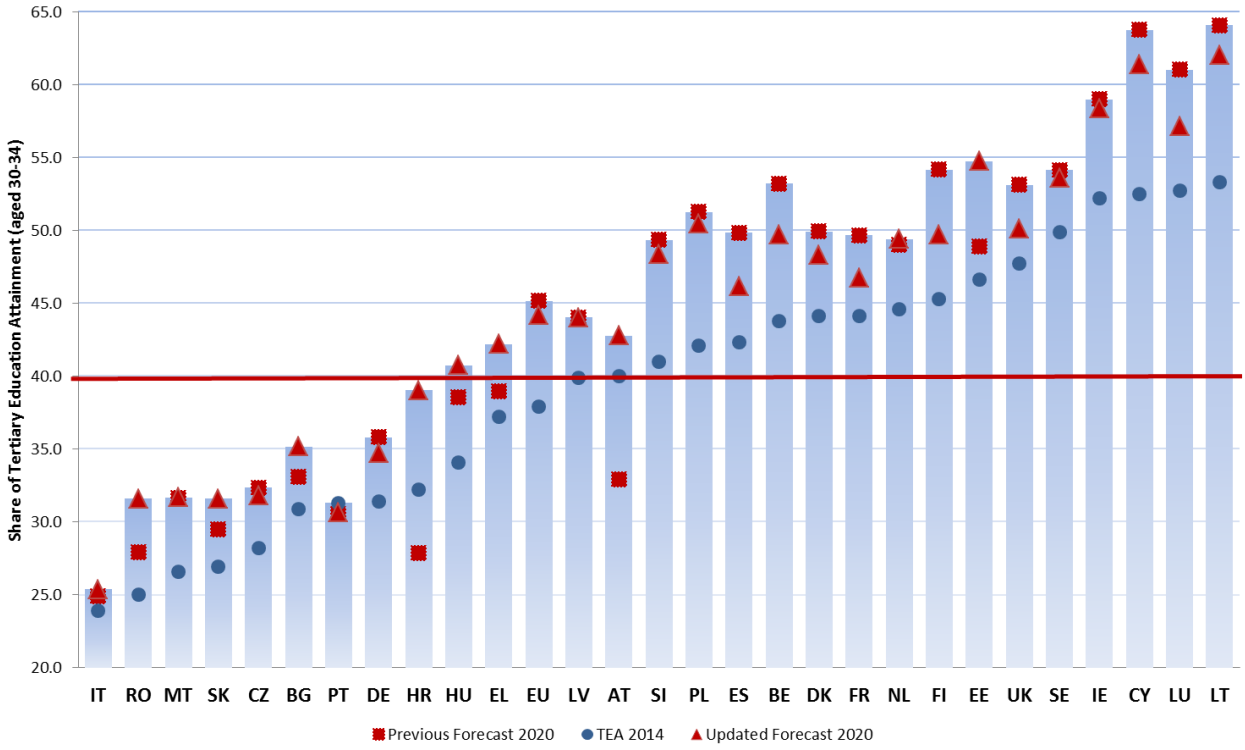
Figure 3 summarizes the outcome of our forecasting exercise for TEA, including country-specific results. While the EU as a whole is indeed likely to reach the 40% target in 2020, this progress masks a vast heterogeneity of countries' progress. As of 2012, the group of countries on the right side of the chart are already beyond the 40% threshold, but the group of countries on the right side substantially

[9] Notice that when running the projections for these variables, projected values begin in 2015. Contrary to the dataset used in Dragomirescu-Gaina et al. (2015), we have real values for the years 2013 and 2014, this may slightly change the results, only for the better as they are real values (we hope).

[10] As indicated by Dragomirescu-Gaina et al. (2015), regarding country-specific forecasts, one major caveat though relates to the importance of country heterogeneity in the context of a future probable convergence process in higher education. If this were to be the case, a country starting, for example, from a low level but with significantly faster than average improvements in education attainment would have an underestimated forecast for TEA based on a model that omits the country-specific trends from equations (2). The first argument against this interpretation is that such model specification did not pass the required autocorrelation tests. By not including country-specific factors in the preferred model specification, extending the model to other countries becomes straightforward.

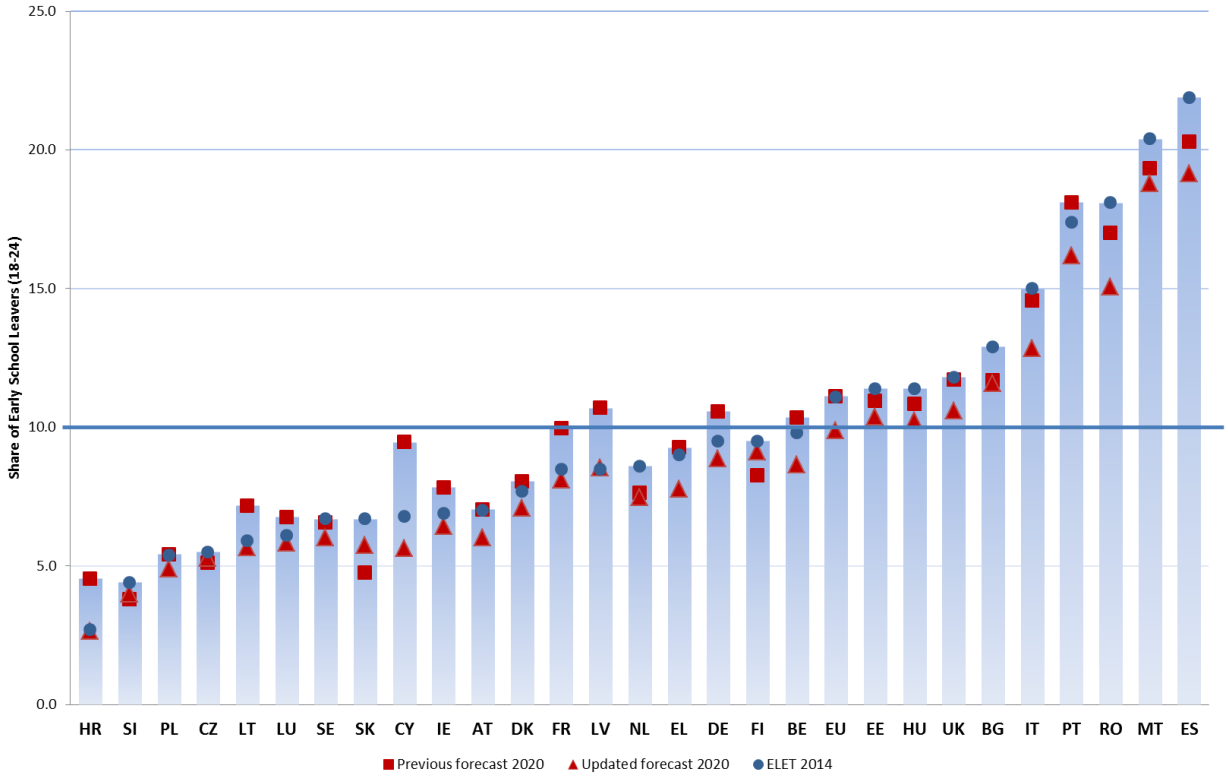
lag behind. Moreover, when looking at the expected progress by the two country groups, it becomes clear that the divergence we observe with respect to 2012 values is also likely to persist in the future. Our calculations show a slower expected progress for those countries currently lagging behind, and a faster expected progress for the high performing countries. In particular, 11 countries are not expected to reach the 40% threshold, namely Italy, Romania, Malta, Slovakia, Czech Republic, Bulgaria, Portugal, Germany, Hungary, Greece and Austria according to the previous forecast. With the updated version, some improvements are observed for Hungary, Greece and Austria. Croatia, which was not included in the former version of the analysis, is also likely to reach this benchmark.

**Figure 3.** Country-specific forecast for 2020 and 2014 share of TEA




Regarding ELET, the picture (see Figure 4) seems equally heterogeneous. Considering that countries must lower down the share of ELET to less than 10%, we can see that countries like Spain, Malta, Romania, Portugal, Italy, Bulgaria and even the UK are very unlikely to reach this benchmark in 2020. Germany, Belgium and Latvia where close to reach the target in the previous forecast and it seems that with the updated version including 2 more years and accurate data they are more likely to reach it.

**Figure 4.** Country-specific forecast for 2020 and 2014 ELET share




Notwithstanding, having worked with a limited data set, the best practice as proposed in the former forecast exercise is to avoid giving point forecasts and instead to highlight the uncertainty as a separate outcome of this type of exercise. Following Tay and Wallis (2000) concerning the production and presentation of conditional density forecasts, all EU28 countries are assessed according to the probability of reaching the targets based on the forecasted 2020 distribution probability.

We will apply the following arbitrary classification:

- *high* probability means that 66% of the expected outcomes would lie below the target in the case of *ELET* / above the target in case of *TEA*.
- *medium* probability corresponds to between 33% and 66% of the expected outcomes.
- *low* probability corresponds to less than 33% of the expected outcomes lying below target in the case of *ELET* / above target in the case of *TEA*.

An overview of the forecasts' results can be observed in Tables 3 and 4 below.

**Table 3.** Qualitative ranking scale for the probability of reaching TEA 2020 target

| | | 2020 probability of reaching TEA target (*100) | | | | |
|---|---|---|---|---|---|---|
| | | Previous forecasts | | | Updated forecasts | |
| | **National target** | **EU target (>=40%)** | **National target** | | **EU target (>=40%)** | **National target** |
| | | **Column 1** | **Column 2** | | **Column 3** | **Column 4** |
| BE | 47 | 1.00 | 0.88 | | 0.99 | 0.73 |
| BG | 36 | 0.04 | 0.22 | | 0.08 | **0.40**[+] |
| CZ | 32 | 0.02 | 0.54 | | 0.01 | 0.47 |
| DK | 40 | 0.98 | 0.98 | | 0.98 | 0.98 |
| FI | 42 | 1.00 | 0.99 | | 0.99 | 0.97 |
| DE | 42 | 0.15 | 0.07 | | 0.05 | 0.02 |
| UK | - | 1.00 | - | | 0.99 | - |
| IE | 60 | 1.00 | 0.44 | | 1.00 | 0.38 |
| IT | 26-27 | 0.00 | 0.34 | | 0.00 | 0.39 |
| NL | 40 | 0.97 | 0.97 | | 0.99 | 0.99 |
| PL | 45 | 0.99 | 0.89 | | 1.00 | 0.90 |
| SE | 40-45 | 1.00 | 1.00 | | 1.00 | 1.00 |
| AT | 38 | 0.03 | 0.09 | | **0.77**[+] | **0.90**[+] |
| CY | 46 | 1.00 | 1.00 | | 1.00 | 1.00 |
| EE | 40 | 0.97 | 0.97 | | 1.00 | 1.00 |
| FR | 50 | 0.98 | 0.47 | | 0.96 | **0.22**[(-)] |
| SK | 40 | 0.00 | 0.00 | | 0.00 | 0.00 |
| SI | 40 | 0.97 | 0.97 | | 0.98 | 0.98 |
| PT | 40 | 0.01 | 0.01 | | 0.00 | 0.00 |
| EL | 32 | 0.40 | 0.96 | | **0.72**[+] | 1.00 |
| ES | 44 | 0.98 | 0.88 | | 0.95 | 0.70 |
| HR | 35.00 | 0.00 | 0.02 | | **0.39**[+] | **0.89**[+] |
| LV | 34 | 0.80 | 0.99 | | 0.85 | 1.00 |
| LT | 48.7 | 1.00 | 0.99 | | 1.00 | 1.00 |
| LU | 66 | 1.00 | 0.27 | | 1.00 | 0.07 |
| MT | 33 | 0.02 | 0.35 | | 0.01 | 0.33[(-)] |
| HU | 30.3 | 0.36 | 0.99 | | 0.58 | 1.00 |
| RO | 26.7 | 0.00 | 0.66 | | 0.00 | 0.97 |
| EU | 40 | 0.87 | 0.87 | | 0.86 | 0.86 |

**Colour coding for the probability of reaching the target**

| | | |
|---|---|---|
| | > 66% | **High** |
| | >33-66% | **Medium** |
| | <=33% | **Low** |

Note: If national targets were set as intervals, they were approximated up to the most conservative value. To interpret results you compare column 1 with column 3 and column 2 with column 4. We calculate the probabilities assuming a normal distribution function.

[+]It means that the situation improves in the updated forecast with respect to the previous one.

[(-)] It means that the situation gets worse in the updated forecast with respect to the previous one.

**Table 4.** Qualitative ranking scale for the probability of reaching ELET 2020 target

| | | 2020 probability of reaching ELET target (*100) | | | | |
|---|---|---|---|---|---|---|
| | | Previous forecats | | | Updated forecast | |
| | **National target** | **EU target (<=10%)** | **National target** | | **EU target (<=10%)** | **National target** |
| | | **Column 1** | **Column 2** | | **Column 3** | **Column 4** |
| BE | 9.5 | 0.45 | 0.38 | | **0.72**[+] | 0.65 |
| BG | 11 | 0.29 | 0.41 | | 0.28 | 0.42 |
| CZ | 5.5 | 0.99 | 0.60 | | 0.99 | 0.57 |
| DK | 10 | 0.78 | 0.78 | | 0.92 | 0.92 |
| FI | 8 | 0.75 | 0.45 | | **0.65**[(-)] | **0.30**[(-)] |
| DE | 10 | 0.42 | 0.42 | | **0.69**[+] | **0.69**[+] |
| UK | | 0.29 | - | | **0.41**[+] | - |
| IE | 8 | 0.81 | 0.53 | | 0.96 | **0.81**[+] |
| IT | 16 | 0.09 | 0.63 | | 0.15 | **0.81**[+] |
| NL | 8 | 0.83 | 0.57 | | 0.88 | 0.61 |
| PL | 4.5 | 0.99 | 0.25 | | 1.00 | **0.37**[+] |
| SE | 10 | 0.93 | 0.93 | | 0.98 | 0.98 |
| AT | 9.5 | 0.90 | 0.86 | | 0.98 | 0.97 |
| CY | 10 | 0.58 | 0.58 | | 0.99 | **0.99**[+] |
| EE | 9.5 | 0.38 | 0.31 | | 0.44 | **0.36**[+] |
| FR | 9.5 | 0.50 | 0.43 | | **0.80**[+] | **0.74**[+] |
| SK | 6 | 1.00 | 0.79 | | 0.99 | **0.57**[(-)] |
| SI | 5 | 1.00 | 0.83 | | 1.00 | 0.82 |
| PT | 10 | 0.02 | 0.02 | | 0.03 | 0.03 |
| EL | 9.7 | 0.61 | 0.56 | | **0.85**[+] | **0.81**[+] |
| ES | 15 | 0.01 | 0.14 | | 0.00 | 0.16 |
| HR | 4.00 | 1.00 | 0.32 | | 1.00 | **0.95**[+] |
| LV | 10 | 0.41 | 0.41 | | **0.74**[+] | **0.74**[+] |
| LT | 9 | 0.87 | 0.78 | | 0.99 | 0.96 |
| LU | 10 | 0.92 | 0.92 | | 0.98 | 0.98 |
| MT | 10 | 0.01 | 0.01 | | 0.01 | 0.01 |
| HU | 10 | 0.39 | 0.39 | | 0.46 | 0.46 |
| RO | 11.3 | 0.03 | 0.07 | | 0.05 | 0.12 |
| EU | 10 | 0.35 | 0.29 | | 0.52 | **0.43**[+] |

**Colour coding for the probability of reaching the target**

| | |
|---|---|
| > 66% | **High** |
| >33-66% | **Medium** |
| <=33% | **Low** |

[+]It means that the situation improves in the updated forecast with respect to the previous one.
[(-)] It means that the situation gets worse in the updated forecast with respect to the previous one. Note: To interpret results you compare column 1 with column 3 and column 2 with column 4. We calculate the probabilities assuming a normal distribution function.

## 5. Discussion and Conclusion

This technical brief updates the forecast of Europe 2020 headline targets on education and training undertaken by Dragomirescu Gaina and Weber (2013). Our results confirm the relatively positive outlook across the European Union regarding the accomplishment of the targets set for both early leaving from education and training (ELET) and tertiary education attainment (TEA). Bulgaria, Austria and Greece have improved either in their national or EU targets for TEA while the number of countries that have improved their probability of reaching the ELET target (national or EU) is much larger (roughly 13 countries). Thus, we can conclude that there seems to be an overall general improvement of those countries lagging behind.

Beyond the policy implications of this exercise, which can be consulted in the previous publication, it is important to refresh and highlight two issues that need to be taken into account:

1) The key role of country expert in making qualitative judgements. The forecasts presented here illustrate only a baseline scenario and could therefore be complemented for each country with expert judgement in order to account for factors not included in our modelling approach. This might change the evaluation presented here for some countries.

2) The limitations related to data availability. This issue hinders a full investigation of the determinants of education decisions. While an update of 2 years does not seem to change the robustness of the specification significantly, we cannot guarantee that these results will hold in the long run i.e. as more data become available.

## REFERENCES


Acemoglu, D. (2002). Technical Change, Inequality, and the Labour Market. Journal of Economic Literature, 40, 7-72.

Acemoglu, D. & Pischke, J. S. (2001). Changes in the wage structure, family income, and children's education. European economic review, 45 (4-6), 890-904.

Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. Review of Economic Studies 58: 277–297.

Asteriou, D., & Agiomirgianakis, G. M. (2001). Human capital and economic growth: time series evidence from Greece. Journal of Policy Modeling, 23(5), 481-489.

Autor, D., Katz, L. F. & Krueger, A. B. (1998). Computing Inequality: Have Computers Changed the Labour Market? The Quarterly Journal of Economics 113 (4): 1169-1213.

Baltagi, B. H. & Wu, P. X. (1999). Unequally spaced panel data regressions with AR(1) disturbances. Econometric Theory, 15: 814–823.

Becker, G. S. & Tomes, N. (1979). An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility. Journal of Political Economy, 87, 1153–89.

Becker, G.S. & Tomes, N. (1986). Human Capital and the Rise and Fall of Families, Journal of Labour Economics 4, no. 3, pt. 2, S1-S39.

Black, S. E., & Devereux, P. J. (2011). Recent developments in intergenerational mobility, in Ashenfelter, O. & Card, D. (eds), Handbook of labour economics, 4, 1487-1541.

Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data model, Journal of Econometrics, 87(1): 115-143.

Cameron, S. V., & Heckman, J. J. (1998). Life cycle schooling and dynamic selection bias: Models and evidence for five cohorts (No. w6385). National Bureau of Economic Research.

Chevalier, A., Denny, K. & McMahon, D. (2003). A Multi-Country Study of Inter-Generational Educational Mobility. In Dolton, P., Asplundh, R. & Erling, B. (eds). Education and Inequality Across Europe. Edward Elgar, London.

Chevalier, A., Harmon, C., O'Sullivan, V. & Walker, I. (2013). The impact of parental income and education on the schooling of their children. IZA Journal of Labor Economics, 2:8.

Dragomirescu-Gaina, C. & Weber, A. (2013). Forecasting the Europe 2020 headline target on education and training. A panel data approach. EUR 26077 EN. Luxembourg: Publications Office of the European Union; JRC83346

Dragomirescu-Gaina, C., Weber, A. and Elia, L. (forthcoming). A fast-forward look at tertiary education attainment in Europe 2020. Journal of Policy Modeling

Hertz, T., Jayasundera, T., Piraino, P., Selcuk, S., Smith, N. & Verashchagina, A. (2007). The Inheritance of Educational Inequality: International Comparisons and Fifty-Year Trends The B.E. Journal of Economic Analysis & Policy, 7(2): 1935-1682.

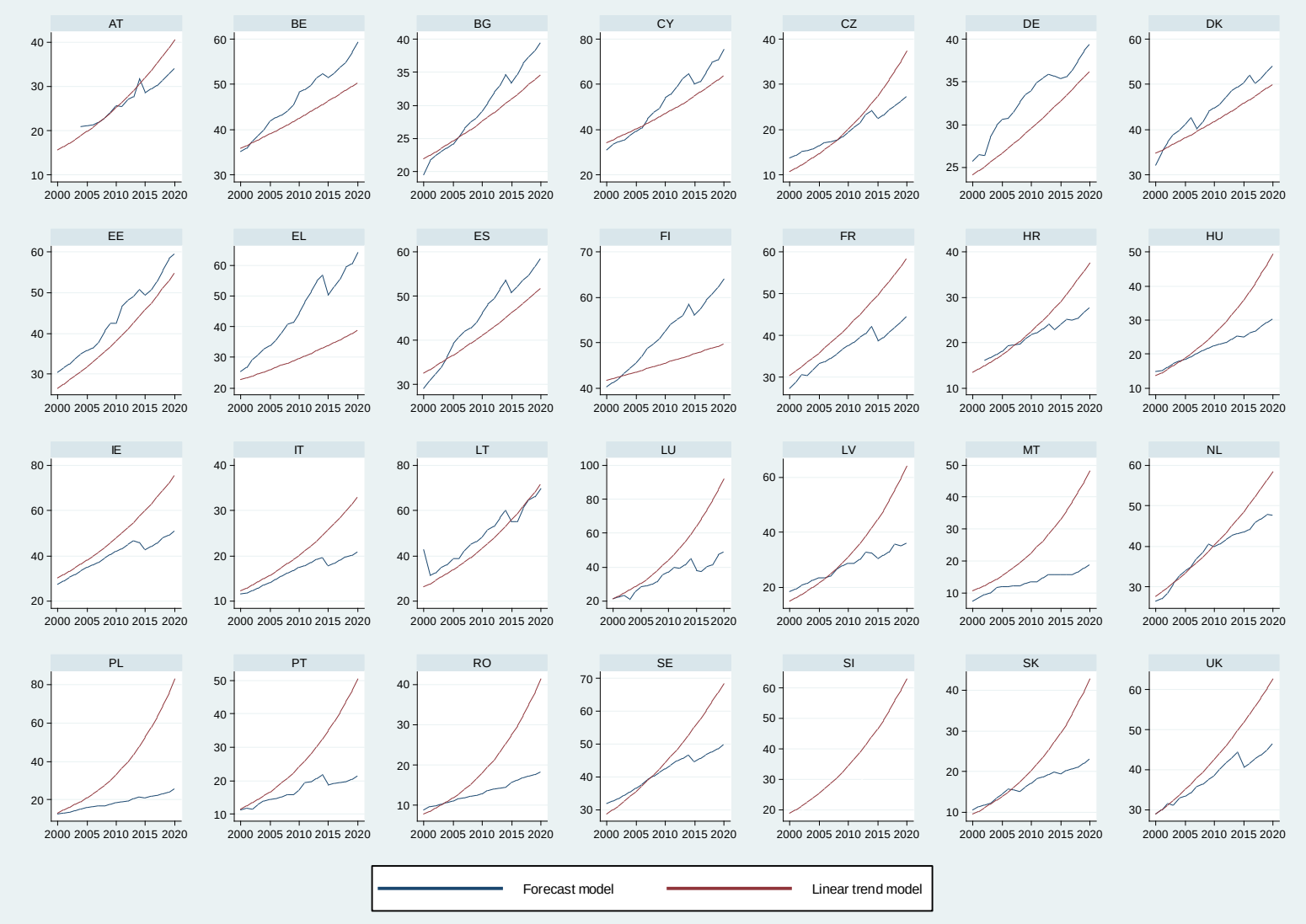
Li, G. (1985). Robust regression. In D. C. Hoaglin, F. Mosteller, and J. W. Tukey, (eds), Exploring Data Tables, Trends, and Shapes, 281–340. New York: Wiley.

Winchester, N., & Greenaway, D. (2007). Rising wage inequality and capital–skill complementarity, Journal of Policy Modeling, 29(1), 41-54.

Wooldridge, J. M. (2002). Econometric Analysis of Cross Section and Panel Data. Cambridge, MA, MIT Press.

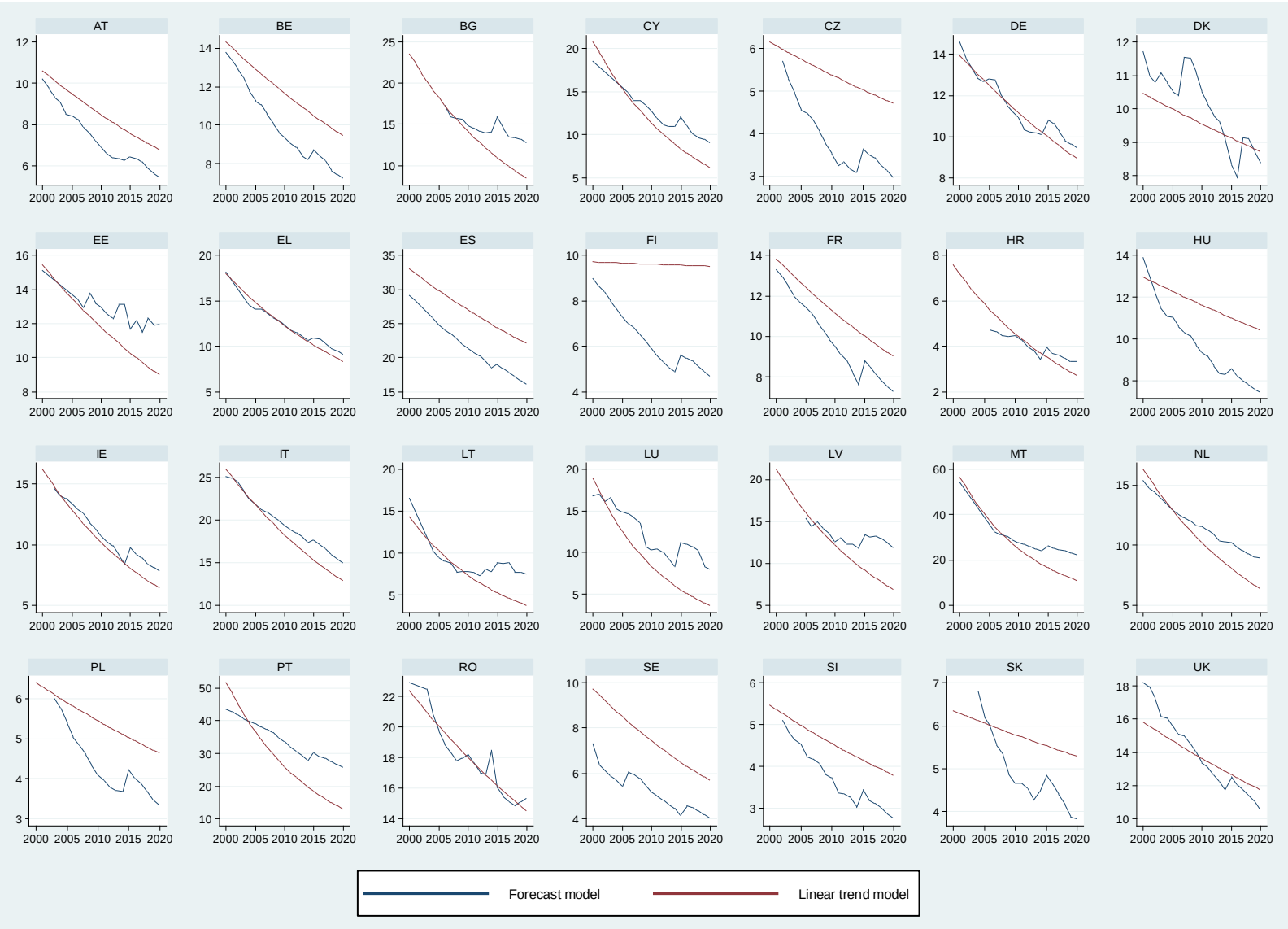
## APPENDIX 1. COUNTRY SPECIFIC FORECASTS

**Figure 1.** Comparison between Forecast model and linear trend model: Country specific regressions for TEA.




Note: Charts in this Figure compare the forecast for the given years (2000-2020) using the econometric model proposed in this technical brief with a simple linear trend model for each country. Please note that the scale of the vertical axis might be different depending on the country.

**Figure 2.** Comparison between Forecast model and linear trend model: Country specific regressions for ELET.




Note: Charts in this Figure compare the forecast for the given years (2000-2020) using the econometric model proposed in this technical brief with a simple linear trend model for each country. Please note that the scale of the vertical axis might be different depending on the country.

European Commission
EUR 26077 EN – Joint Research Centre –
Deputy Director-General Office, Econometrics and Applied Statistics Unit

Title: Education and youth labour market outcomes: the added value of VET

Author(s): Esperanza Vera-Toscano

Luxembourg: Publications Office of the European Union

2015 – 22 pp. – 21.0 x 29.7 cm

EUR – Scientific and Technical Research series – ISSN 1831-9424

ISBN 978-92-79-51831-7

doi:10.2788/05658

Abstract
This technical brief builds on the work undertaken by Dragomirescu-Gaina and Weber (2013), aiming to provide an update to the question whether Europe as a whole and each of its Member States will reach the twofold Europe 2020 target on early leavers from education and training (ELET) and tertiary education attainment (TEA) by 2020.

LB-1A-26077-EN-N

## JRC Mission

As the Commission's in-house science service, the Joint Research Centre's mission is to provide EU policies with independent, evidence-based scientific and technical support throughout the whole policy cycle.

Working in close cooperation with policy Directorates-General, the JRC addresses key societal challenges while stimulating innovation through developing new methods, tools and standards, and sharing its know-how with the Member States, the scientific community and international partners.

*Serving society*
*Stimulating innovation*
*Supporting legislation*

Publications Office

doi:10.2788/05658

ISBN 978-92-79-51831-7