Forecasting the Europe 2020 headline target on education and training: A panel data approach

Dragomirescu-Gaina, Catalin and Weber, Anke

European Commission, Joint Research Centre, Institute for the Protection and Security of the Citizen, European Commission, Joint Research Centre, Institute for the Protection and Security of the Citizen

2013

JRC SCIENTIFIC AND POLICY REPORTS

Forecasting the Europe 2020 headline target on education and training

A panel data approach

Catalin Dragomirescu-Gaina
Anke Weber

2013

Report EUR 26077 EN

Joint Research Centre

European Commission
Joint Research Centre
Institute for the Protection and Security of the Citizen (IPSC)

Contact information
Catalin Dragomirescu-Gaina
Address: Joint Research Centre, Via Enrico Fermi 2749, TP 361, 21027 Ispra (VA), Italy
E-mail: catalin.dragomirescu-gaina@jrc.ec.europa.eu
Tel.: +39 332 78 9447
Fax: +39 332 78 5733

Contact information
Anke Weber
Address: Joint Research Centre, Via Enrico Fermi 2749, TP 361, 21027 Ispra (VA), Italy
E-mail: anke.weber@jrc.ec.europa.eu
Tel.: +39 332 78 5821
Fax: +39 332 78 5733

http://ipsc.jrc.ec.europa.eu/
http://www.jrc.ec.europa.eu/

JRC83346

EUR 26077 EN

ISBN 978-92-79-32509-0 (pdf)
ISBN 978-92-79-32510-6 (print)

ISSN 1831-9424 (online)
ISSN 1018-5593 (print)

doi:10.2788/92332

Luxembourg: Publications Office of the European Union, 2013

Printed in 2013

Forecasting the Europe 2020 headline target on education and training – a panel data approach* –

Catalin Dragomirescu-Gaina and Anke Weber

Executive summary

Education is a major pillar of the Europe 2020 strategy[1]. Its long-term impact on economic growth and productivity is widely recognised. However, under fiscal constraints, challenging demographic developments and alarmingly high levels of youth unemployment, education is becoming decisive for the future of European economic progress. In this context, policy and reform agendas require early indicators and a suitable monitoring framework to guide them.

This analysis aims at proposing simple econometric models that can be used to forecast the twofold Europe 2020 headline target on *early leavers from education and training* and *tertiary education attainment*. These are at the same time two out of seven benchmarks in the strategic framework for European cooperation in education and training[2] ("ET 2020").

The models are built on the theoretical framework of human capital and then estimated in a panel setting to better deal with a limited dataset. According to our findings, *early school leavers*[3] are sensitive to employment opportunities and are highly dependent on the educational attainment of the adult population in the age groups corresponding to the parental cohort. *Tertiary education attainment* depends on labour productivity as a proxy for expected wage differentials, and adults' educational level as a proxy for family background. Under given methodological constraints, the benchmarks were designed to reflect the consequences of education decisions taken previously. Therefore, by looking back at the time period of enrolment and graduation, our approach could be seen as an attempt to identify the factors that shape the education decisions of young individuals.

We construct the forecasts under simple but realistic assumptions about the expected level of adults' educational attainment, given the determinants of schooling decisions uncovered by our empirical analysis. The forecasts tell us how early school leaving and tertiary education attainment are likely to develop over the next years *if nothing changes in terms of policy measures*. This very strong assumption provides scope for policy action especially for those countries where the expected developments of model's determinants are not enough to foresee a positive outcome.

Uncertainty is an integral part of our exercise, so we present the forecasts as confidence intervals allowing us to construct a qualitative evaluation of the probability that each Member State would reach the targets in education and training by 2020. Our results paint an optimistic outlook for the majority of EU27 countries, more precisely for 16 Member States in case of *early school leavers* and for 15 in case of *tertiary education attainment*. Some countries still have room to improve their outcomes, which under the current assumptions are not yet satisfactory, while very few would need sustained efforts and active policy initiatives to increase their odds of attaining the targets.

* We would like to thank Luca Pappalardo, Leandro Elia, Violeta Piculescu and Stan van Alphen for helpful comments and suggestions.

[1] See COM(2010) 2020 final.

[2] Adopted by the Council in May 2009 (2009/C 119/02).

[3] The terms *early school leavers, early school leaving* and *early leavers from education and training* are used interchangeably in this paper.

We highlight family background as a policy lever and suggest the potential benefits with regard to policies designed to encourage lifelong learning, remove barriers to participation in training and promote outreach programs towards the disadvantaged. Meanwhile, policies that provide adequate counselling about the long term benefits of education and related job prospects, or improve resource allocation in the labour market and reduce skills mismatches, could develop the appropriate economic incentives for better education attainment.

1. Introduction

The standard theoretical framework[4] for studying education attainment draws on human capital models of life cycle earnings and education investment developed by Becker (1964, 1975), Ben Porath (1967), and Mincer (1974). Human capital represents the stock of knowledge, skills or characteristics, either *innate* or *acquired*, which contributes to workers' productivity. Education – as a main component of human capital[5] – usually comes through formal schooling, a type of investment with upfront costs and delayed benefits. Individuals value employment and earning prospects against opportunity costs of not entering the labour market earlier and other schooling related expenses (e.g. tuition). According to a simple model of human capital accumulation, an optimal schooling decision would depend mainly on expected lifetime earnings from labour[6].

This study builds on the theoretical framework outlined above in order to construct econometric models capable of explaining the evolution of early school leaving[7] and tertiary education attainment[8] *over time* and *across Member States*. To this end we are going to use a set of determinants drawn from theory that describes labour market and economic interactions with schooling decisions. Our task is complicated by the multitude of effects and causality relations between education and economic developments, but we are relying on rich empirical evidence to guide us through.

As already mentioned, the focus of the study will be on explaining the *improvement*[9] over time in these two benchmark indicators. We adopt a modelling strategy specific to a panel setting, meaning that we would more likely observe common trends and similarities. This means we are going to explain what drives the change in education attainment at the country level by using a common set of determinants. For example, country specific cultural factors that could affect education motivation and traditions towards learning and studying will not be distinguishable within our models. This is because these factors are only changing very slowly over time, so they could not significantly affect the change in *early school leaving* and *tertiary education attainment*. On the contrary, youth education attainment would be changing over time in relation to adult education attainment, encompassing parental influences and family bonds.

An empirical model that would provide the best explanatory value as regards the two benchmark indicators would not necessarily be the most suitable from a forecasting perspective. While using a multitude of explanatory variables could improve the overall data fit, it could also

[4] The leading alternative views education as a signal for individuals' unobserved innate abilities. This approach would usually be taken by cross sectional studies, using mainly data from surveys where heterogeneity is better accounted for. This approach would not be taken here, mainly because the dataset we are using and the purpose of the analysis are different, as will be explained later.

[5] There are other ways to accumulate human capital e.g. on-the-job training which will not be considered here.

[6] Other important determinants emphasized in these models relate to discount factors, life expectancy and mortality risk, skill premium etc. See Ben Porath (1974), Becker (1975), Heckman (1976) and others.

[7] *Early leavers from education and training* are persons aged 18 to 24 fulfilling the following two conditions: *(1)* the highest level of education or training attained is ISCED 0, 1, 2 or 3c short and *(2)* no education or training has been received in the four weeks preceding the survey. While there is a strong variety in national targets set by the Member States, the EU-wide target is to have less than 10% early school leavers by 2020.

[8] *Tertiary education attainment* refers to persons aged 30 to 34 years who have successfully completed university or university-like (tertiary-level) education with an education level of 5-6 following the International Standard Classification of Education (ISCED) 1997. While there is a strong variety in national targets set by the Member States, the EU-wide target is to have at least 40% tertiary education attainment by 2020.

[9] Improvement here is a generic term and it refers to changes in both directions.

affect the stability over time of the model coefficients and inappropriately lower the estimated uncertainty which is an important component of the forecast itself. By over-fitting the data, we run the risk of being too confident when anticipating the future based on the developments observed over a short time interval. Therefore, we select the best specification based on comparison of historical data with models' predictions for *out of sample* and coefficients' stability over time.

There are mainly two distinct approaches identifiable in the literature to forecast education attainment, depending upon both the purpose and the methods involved.

The first approach would be the most common one and builds on standard assumptions about, *inter alia*, demographic developments plus enrolment and completion rates in education. It could be adapted for long and very long term forecasting exercises that can go as far as 50 years ahead. Some international organisations, such as UNESCO [10] and the World Bank [11] are developing and using these **cohort-based models**, which allow for the inclusion of countries at different stages of development or with different education systems.

The second approach addresses education from a labour economics perspective, which regards skills and education attainment as close equivalents. The approach examines the interaction between demand for and supply of education, offering insights about expected skill mismatch bottlenecks. There are a number of government agencies/bodies – such as CEDEFOP in Europe[12] and the Bureau of Labour Statistics[13] in the United States – providing occupational and education forecasts that usually span over a 5-to-10 years horizon. The methods used are more complex and rely on a macroeconomic model able to foresee the changing economic structure by incorporating business cycle dynamics. Some assumptions regarding the required education level by occupation within each industry are also required to arrive at estimates of job creation.

This paper takes a different approach, more in line with the evidence provided by the empirical literature on education attainment. We develop **econometric models** that emphasize family background together with employment and earning prospects as the main factors driving education attainment over time. As the benchmarks are designed to measure education attainment *after* individual schooling decisions have been taken[14], we use past values of the driving factors meant to capture the labour market conditions or economic context at the relevant point in time. This way we can ensure better links with the enrolment and graduation decision moments and mitigate possible reverse causality issues.

Later in this paper we will compare the forecasts obtained using the econometric model for tertiary education attainment against those derived from the birth cohort approach. Although both methods build on different assumptions, the comparison provides further evidence on the relevance of our econometric modelling approach in explaining education attainment.

The study is structured as follows. Section 2 describes the theoretical model and provides a survey on the relevant literature on education attainment. Section 3 summarizes the statistical data and presents the econometric approach. Section 4 highlights the results of this forecasting

[10] See Lutz and Scherbov (2006) for details about the approach and the software used to generate forecasts.

[11] The World Bank coordinates some different projects in this area and has a full set of data, models and projections (for indicators such as education attainment, enrollment rates, gender parity etc) available at http://go.worldbank.org/DKACUHA0D0

[12] Except CEDEFOP, in Europe there are some other universities, research centers and other government institutions involved in occupational forecast and skills mismatch analysis. Examples include: Research Centre for Education and the Labour Market (ROA) in the Netherlands, Economic and Social Research Institute (ESRI) in Ireland, Statistics Norway etc. Please refer to Campos et al. (1999) for a nice survey of the methods used in several European countries and to Bjørnstad and Gjelvsik (2011) for Norway.

[13] For more information see B.L.S. Occupational Outlook Handbook, 2012-13 Edition, available at http://www.bls.gov/ooh/home.htm

[14] This assertion would mainly disregard the "not in training" share of individuals included in the early school leavers benchmark. A quick look at the statistical data would show that this component does not have a significant share although it can rise in the future given an increasing government recourse to active labour market policies on the back of high youth unemployment in various European countries.

exercise and discuss some policy relevant issues. Section 5 presents the birth-cohort model for the tertiary educational attainment benchmark and the underlying assumptions. Section 6 concludes, outlining the potential role for policy action and some possible future research directions.

2. Theoretical background

Ben Porath (1967) pioneered a theoretical model in which individual investment decisions are based on maximising the present value of lifetime earnings from labour. Under certain simplifying assumptions[15], he proved that the optimal schooling decision depends on mortality risk, discount rates and expected earnings from labour. Lower mortality risk and lower discount (or interest) rates increases schooling, whereas lower future labour income decreases it. This simple theoretical framework has been a fertile ground for subsequent empirical and theoretical work. For example, it provided the basis for Mincer (1974) to develop his equation explaining wages as a function of years of schooling, which represents the keystone of much empirical literature.

Adding borrowing constraints into the analysis, Becker (1975) and then Becker and Tomes (1979, 1986) highlight the positive relationship between family income and children's education attainment[16]. There is now widespread consensus in the empirical literature concerning innate factors and family bonds in explaining the *innate* abilities of individuals and accounting for human capital differences[17]. Wealthy parents have more resources available for their children's education, thus relaxing the financial constraints that are binding, especially at higher education levels.

But, apart from family influences, the two benchmark indicators are ultimately driven by the education choices made by young individuals in the age groups 18 to 24 (in the case of *early school leavers*) and 30 to 34 (in the case of *tertiary education attainment*). For a better characterisation of their education decisions, we need to understand what shapes their expectations about *employment* and *earning prospects*, the main factor behind schooling decisions according to theory.

Early school leavers 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. Pissarides (1981) was among the first to observe the cyclical component of dropout rates and the myopic reaction to cyclical swings in economic activity. He studied the staying-on rate[18] for 16-year-olds in the UK and found a positive relation between adult unemployment and enrolment in post-secondary education using time series analysis[19]. Whitfield and Wilson (1991) studied the same problem but used newly developed co-integration techniques[20]. Mattila (1982) and Betts and McFarland (1995) reached the same conclusions for the United States. Other studies such as Grubb (1988), Micklewright et al. (1990) were less conclusive about the link between dropout rates and

[15] Among the assumptions of the original model we mention: lack of leisure in the utility function, timeless perspective implying exogenously fixed wages and continuous labour participation after schooling etc. Please pay attention that in the original framework schooling would refer to the number of years of schooling *per se*, while the ET2020 headline targets would more likely refer to having obtained credentials (i.e. diploma or degrees) after fulfilling the final year of certain educational segments.

[16] See for example Jacoby (1994), Keane and Wolpin (2001) etc

[17] See Haveman and Wolfe (1995) for a nice survey of some earlier papers. More recent studies include Ermisch and Francesconi (2000), Cameron and Heckman (2001) , Tieben and Wolbers (2010), Mocetti (2010) etc.

[18] Staying-on rate could be interpreted as the *inverse* of early school leavers' share, although the age groups are different.

[19] Pissarides used a logistic transformation in order to ensure that his estimated probabilities stayed in the [0,1] interval, working therefore with participation rates or data in *levels*. We tried the same approach in the beginning but given our forecasting purposes and the potential non-stationarity issues related to education attainment indicators we had to differentiate the data to obtain stationarity.

[20] They used a larger data set that spans about 30 years (1956 - 1985).

unemployment or its proxies, but pointed to institutional rigidities as potential explanations. 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.

We specify the model for *early school leavers* as a function of parents' education together with employment prospects. We therefore expose the link between education decisions and employment opportunities by using different indicators of economic cycle that also capture labour market conditions. Appropriate counselling about the long term benefits of education and related employment prospects would improve schooling attainment for very young individuals.

Highly educated individuals are seen as complements to physical capital in the production process, helping innovation and competitiveness. They have longer time horizons, high and stable employment rates that vary little with the business cycle (see the figure in Annex 4) and higher earning profiles. The (expected) wage or the so called *skill premium* would be more important for their education investment decisions than a job offer *per se*. Autor et al. (1998) and Acemoglu (2000) among others found that technology shifts over recent decades have favoured skilled workers. They noticed that despite a pronounced increase in the share of high educated workers entering the labour market, the wage premium has not declined but instead increased further. Their observation suggests the presence of a positive feedback loop between skills and wages on the back of productivity upgrades. Restuccia and Vandenbrouche (2013) have recently proposed a model explaining education attainment differences across countries and over time, where labour productivity and life expectancy play significant roles. Bils and Klenow (2000) question the explanation for the strong empirical relation between education and growth given in Barro (1991), Barro and Sala-i-Martin(1995) and many others. They provide a discussion about the reverse causality channel where expected growth might affect schooling. The primary motive is that expected growth would reduce discount rates (or real interest rates), increasing demand for schooling. They remark that: the more growth is foreseen, the bigger its effect on schooling and the larger the role of reverse causality. Buchinsky and Leslie (2010) also insist on correctly anticipating wage differentials in the education decision process.

Our specification for *tertiary education attainment* includes family education background and different productivity measures to proxy for income prospects. We therefore highlight the transmission channel between productivity gains and real wage increases as a main driver of education decisions. Improving this transmission mechanism would reduce uncertainty about expected wage differential (or *skill premium*) and smooth the education decision process. This in turn would offer the right incentives for young individuals to enrol in universities and graduate with the highest possible education level in anticipation of a higher labour income stream.

3. Data and empirical approach

All the historical data used in this report has been taken from the European Statistical Office (Eurostat)[21]. 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[22], which represents a drawback for any empirical analysis. Table 1 below provides some details about data availability for both benchmark indicators at EU27 level.

[21] All data have been downloaded up to April 25th 2013.
[22] Data labeled by Eurostat with a (b) handle.

Table 1

Headline indicator	Data range	Number of observations (only EU27)		Comments
		Average per country	Minimum	
Early school leavers	*1992 – 2012*	*16.2*	*11*	*Most of the breaks appear more in 2003 but also in 1999, 2004, 2005* *Missing values for some member states*
Tertiary education attainment	*2000 – 2012*	*12.9*	*9*	*Most of the breaks appear around 2003*

Data on the education attainment of the adult population, which is used here as a proxy for *parents' education*, is also available up to 2012 from Eurostat. We extend the data up to 2020 using a simple extrapolation method that builds on very simple assumptions. More details can be found in Annex 5.

Data on the economic indicators needed in various econometric specifications illustrated in Annex 3 were mainly drawn from the AMECO database maintained by DG ECFIN. The dataset includes both the historical values of the indicators and the official European Commission macroeconomic forecasts[23] for 2013 and 2014.

We take a panel approach in order to mitigate the disadvantages associated with having a limited dataset available for the two benchmarks such as short time series, methodological breaks and missing data for some of the countries[24]. A panel approach would therefore allow us to maximize the information contained in the data.

We opt for an econometric specification in first differences[25] that explains the *improvement* in the Europe 2020 headline targets over time and across countries. Any country specific constant factors, whether institutional, cultural or other, are therefore left aside by this transformation of the data. The approach would mitigate most of the concerns associated with time series stationarity and residual autocorrelation but could, however, miss some of the information contained the original dataset[26]. A standard choice in a panel setting would be to allow for country specific dynamics to be summarized by some specific variables usually denoted as *country dummies*. Although we include them in some alternative specifications, they will not appear in the final ones on which the current forecasts have been based. At least three reasons can be mentioned for our choice: *(i)* first differentiating the data has already purged country specific differences, *(ii)* country heterogeneity[27] seemed to be well captured by the main model's determinants and *(iii)* specifications including country dummies performed worse compared to others when judged according to our model selection criteria (see below).

[23] The forecasts correspond to the "European economic forecast – winter 2013" available as of April 25th 2013 at http://ec.europa.eu/economy_finance/publications/european_economy/forecasts_en.htm

[24] Country specific econometric models for each member state would be hard to imagine as long as there are countries with as many as 3 breaks within a span of 12 years, especially in the case of early school leavers.

[25] We took the differences of the *log* data. This approach would also weaken the argument for using fixed effects, as long as any country specificity would be difference out in the transformed data. All our specifications were estimated with the STATA statistical software.

[26] In case that some variables would share a common stochastic trend in the long run (i.e. would be *cointegrated;* possible candidates in our case would be children education attainment and parents' education attainment) this property could be lost after taking first differences. Nevertheless, this assumption is hard to test using the current data set, as unit root tests lack power in short time series data sets and in the presence of structural breaks.

[27] *Country dummies* explain less than 10% of the variance for early school leavers and less than 15% for tertiary education attainment in a wide range of specifications.

Drawing on the theoretical and empirical literature outlined in section 2, we explored various econometric specifications for the two benchmarks. We avoid over-fitting the models and keep them quite simple given the: *(i)* limited data set available, *(ii)* need for robustness and stability of the model coefficients over time and *(iii)* need for simplicity from a policy making perspective. Simplicity has nevertheless one main drawback, namely the *omitted variables* problem (which is relevant here as we do not intend to build a complete structural model of education attainment). Most of the problems related to autocorrelation have already been mitigated to some extent in our first-difference approach. But given that both benchmarks include consecutive population cohorts, autocorrelation could still remain at higher order lags. For this reason we select models that pass most of specification tests proposed below.

We arrive at the best model specification for each of the two benchmarks following the five steps listed below. We proceed as follows:

a) We select the appropriate lag structure of the model determinants based on Akaike (AIC) and Bayesian (BIC) information criteria;

b) We use both balanced and unbalanced panels depending on data availability of the regressors included. We tried to retain as many countries as possible out of 27 Member States, but some had to be dropped when a zero weight was assigned to them in robust regression estimation[28].

c) We keep only those specifications that passed statistical tests for lack of residual autocorrelation[29] according to tests proposed by Arelano-Bond (1991), Baltagi-Wu (1999) and Wooldridge (2002, 282–283) - Drukker (2003);

d) We check the robustness of our specifications by observing the stability of the coefficients over time when varying the estimation sample. This means that successive re-estimations of a model by adding more recent observations should not significantly change its coefficients.

e) We select the best specification according to *out-of sample* root mean square error[30] (RMSE) as a final criterion computed over a 1-to-4 year horizon.

To better articulate our empirical approach, we gain insights from the available list of sub-indicators[31] designed to monitor the twofold Europe 2020 headline target in education and training. Among these, a key role is assigned to female education attainment as a proxy for family influences and employment rate differentials by education levels as a proxy for employment and earnings prospects. However, we did not restrict our choice to this list of sub-indicators and instead searched for proxies of the theoretical determinants within a larger pool of indicators. The gap between theoretical models of human capital and their empirical counterparts is known and well understood. A large number of potential regressors and poor data availability generally restrict the empirical specifications that can be examined and the methods that can be employed.

A common determinant for both benchmarks is the proxy for family influences meant to capture borrowing constraints as referred to in the theoretical model of human capital accumulation. The empirical literature mentions several indicators but we prefer *parental education* given our

[28] This means that the country excluded is considered an outlier in a robust regression estimation.

[29] Residual autocorrelation would biases the OLS estimates.

[30] Although some of the specifications include only a sub-set of the countries (due to data limitations), we always compute RMSE statistics taking into account forecast accuracy for all 27 member states.

[31] A list of sub-indicators is used to monitor Member States' progress towards the Europe 2020 twofold headline target in education and training according to the methodology of the Joint Assessment Framework (JAF), which was developed by DG EMPL and adapted to the field of education and training by JRC-CRELL for DG EAC. See the JRC-CRELL report by Badescu et al. (2012).

subject matter and because it is a good indicator of parental income[32], job tenure, socio-economic status and other family characteristics that could shape education decisions. In fact, we use *adult population education attainment,* split by gender, and then select the age groups in order to match as closely as possible a typical parental relation[33] between adult cohorts and youth cohorts. In simple terms, we follow over the years, the educational attainment of children belonging to an *average family* constructed using the age brackets outlined in Table 2 below (i.e. we do not study the *same family*, but the *same age brackets*).

Table 2

EU2020 headline target		Proxy for parental education attainment used in the econometric specification		Comments
	Age group for benchmark		Age group for parents	
Early school leavers	*18-24*	*Share of females with at most lower secondary education attainment* *Share of males with at most lower secondary education attainment*	*35-44* *45-54*	*Various alternatives were tried by varying the gender or the age group in different specifications, but found less robust or with lower explanatory power.*
Tertiary education attainment	*30-34*	*Share of adults with tertiary education attainment*	*55-64*	

3.1. Early leavers from education and training

Derived from the approach proposed by Pissarides[34] (1981), our preferred econometric specification uses total unemployment rate as a proxy for *employment prospects*, capturing labour market conditions and business cycle dynamics. By separately including adults' education as a proxy for family background and borrowing constraints, we can interpret[35] the unemployment rate as a summary of the available set of opportunities outside the education system.

Alternative specifications of the model involve various other proxies such as output gap or unemployment by skills (including various unemployment skill differentials) with similar implications. For *early school leavers* we estimate specifications with the general form given by equation below:

$$\Delta \log(\text{early school leavers})_t =$$

$$= \eta + \beta * \Delta \log(\text{parents' education})_t + \varphi * \Delta \log(\text{employment prospects})_{t-k} + \varepsilon_t \qquad (1)$$

[32] We have nevertheless tried including some other proxies of family or parents' income such as GDP per capita but with less success.

[33] The structure of the LFS data 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.

[34] For related approaches see Mattila (1982), Rice (1987), Whitfield and Wilson (1991), Betts and McFarland (1995), Petrongolo and Segundo (2002) among others.

[35] Without including adults' education, the alternative interpretation would have been that unemployment was a proxy for parents' financial resources available to invest in children's education, in which case a reverse relationship should have been observed: higher total unemployment would lead to lower children education via lower parental financial resources.

where η, β and φ are model coefficients, Δ is the first difference operator, k represents the empirically estimated lag length expressed in number of years and ε is the error term.

Unemployment swings are strongly associated with business cycles. According to our empirical specifications (alternative specifications could be found in Annex 3, Table A.2), increasing unemployment rates and economic downturns[36] would, *ceteris paribus*, improve education attainment by lowering the rate of early school leavers. On the other hand, decreasing unemployment or economic booms have a *perverse* impact on education attainment. These results go in line with the empirical observation about the myopic behaviour of early school leavers to economic cycles. Our analysis therefore demonstrates the importance of early intervention and appropriate counselling about the long term benefits of education and related job prospects for young individuals.

From a policy making perspective, our results could suggest an undesirable trade-off between employment and education. This doesn't have to be necessarily true. Our alternative specifications better illustrate the importance of compositional effects and skill mismatches and suggest that policy actions could be adapted according to expected labour market developments. A lower or decreasing unemployment rate could come as a result of *(i)* declining unemployment for unskilled workers or *(ii)* increasing employment of medium-skilled workers[37], but the implications would be quite different for education attainment.

We arrived at the specification displayed in Table 3 below by following the five steps listed above (at the beginning of section 3). The lag specification of the unemployment rate was not imposed *a-priori* but selected empirically. Looking at the age groups included in the benchmark i.e. 18-24 and the estimated lag length, we can interpret the observed overlap as an empirical description of the education decision problem when it comes to enrolment in higher secondary education.

Table 3

Δ log(*early school leavers*)	
Δ log(*share of females' with low education, age 35-44*)	0.13***
	(0.05)
Δ log(*share of males' with low education, age 45-54*)	0.23***
	(0.07)
Δ log(*total unemployment rate*), lag 6th	-0.07**
	(0.03)
Constant	-1.82***
	(0.55)
Observations	276
R^2	0.137
No. of countries	19
Estimation sample	1992-2012
Year dummies†	yes
Country dummies	no

Standard errors in parentheses
* p < 0.10, ** p < 0.05, *** p < 0.01
† Only year dummies before 2000 were included to counter the unbalanced panel specification

[36] We refer here to the specification including the output gap, the standard measure for the business cycle.

[37] According to Eurostat, the employment rate for 15-64 year old individuals at the EU27 level for medium skilled group (ISCED 3-4) has average 68.8% between 2003-2012 while the employment rate for unskilled group (ISCED 0-2) has averaged 46.7% over the same period. The latest available data for the last quarter of 2012 shows a 44% figure for the low skilled group and 67.9% for the medium skilled group.

3.2. Tertiary education attainment

Autor et al. (1998) and Acemoglu (2000) provide a demand-supply framework for studying wage differentials according to skill levels. They advocate that over recent decades, productivity changes have been more biased towards skilled workers, allowing them to enjoy higher relative wages.

The *tertiary education attainment* benchmark is supposed to cover those individuals choosing between upper secondary and tertiary education. Here we follow the idea that *income prospects* are more relevant for their schooling decisions than *employment prospects* as in the case of early school leavers. The estimated equation for *tertiary education attainment* has the following general formulation:

$$\Delta\ \log(\text{tertiary education attainment})_t =$$

$$= \eta + \beta{*}\Delta \log(\text{parents' education})_t + \varphi{*}\Delta \log(\text{income prospects})_{t\text{-}k} + \varepsilon_t \qquad (2)$$

where η, β and φ are model coefficients, Δ is the first difference operator, k represents the lag length in number of years estimated empirically and ε is the error term.

In explaining *tertiary education attainment* we use labour productivity growth[38] as a proxy for income prospects. An increase in labour productivity would trigger more schooling according to our model. The relationship between productivity and wages is not direct or immediate but depends on labour market institutions, wage rigidities, union bargaining power, contractual arrangements and other factors. As a consequence of our analysis we can argue that policy reforms addressing labour market institutions could have important second order long term effects on education attainment.

We arrived at the specification displayed in Table 4 below by following the five steps listed above in section 3. The lag specification was not imposed *a-priori* but again selected empirically. By subtracting the estimated lag length from the age corresponding to the cohorts included in the benchmark calculation i.e. 30-34, we can observe an interesting overlap with the decision time about enrolment. We believe that most of the reverse causality issues would be mitigated by the lag length use in this specification.

Table 4

Δ log(tertiary education attainment)	
Δ log(share of adults with high education, age 55-64)	*0.34****
	(0.06)
Δ log(labour productivity), lag 13th	*0.58***
	(0.23)
constant	*0.48*
	(0.60)
Observations	*144*
R2	*0.322*
No. of countries	*12*
Estimation period	*2001-2012*
Year dummies†	*yes*
Country dummies	*no*

Standard errors in parentheses
* p < 0.10, ** p < 0.05, *** p < 0.01
† Only year dummies for 2003 and 2004 were included to account for breaks in the data

[38] We also tried *total factor productivity* growth (or TFP) with similar results (see Annex 3, Table A.3.2). The 11 years lag of employment growth was a better proxy in terms of in-sample fitting and out-of sample forecasting, but the models' coefficients lacked stability over time in this case.

4. Results of the forecasting exercise

This section details the steps taken to produce country specific forecasts for *early school leavers* and *tertiary education attainment* up to 2020 using the econometric models developed in the previous section. We proceed as follows:

(i) Firstly, we construct country specific projections for all our exogenous variables. Details about how the projections for parental education from 2013 up to 2020 were constructed can be found in Annex 5. For all the other economic indicators we rely on the macroeconomic forecasts provided by the European Commission. The selected specification for *early school leavers* presented in section 3.1 includes a 6 years lag for unemployment which allows us to construct the 2020 forecast without any additional assumption. The econometric specification for *tertiary education attainment* illustrated din section 3.2 includes a 13 years lag for labour productivity, therefore requiring only available data to compute the 2020 forecast.

(ii) Secondly, we compute the expected change in both benchmarks up to 2020 based on the estimated models but conditional on the future path of the model's exogenous regressors (available after the previous step). Under normality assumptions, we derive the upper and lower bounds for given different probability levels[39] in order to better illustrate the uncertainty inherent in the forecasts.

(iii) Thirdly, we compute country specific forecasts for both benchmarks based on data obtained at the previous step and subsequently calculate the probability of reaching the EU and the national targets by 2020[40].

Because the forecasts produced under the current approach are ***conditional forecasts*** we must lay down the assumptions under which they are valid. We have assumed:

- 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 2012. 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 assumptions could allow us to work under different alternative scenarios for parental education, other than the one illustrated in Annex 5.
- 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.

[39] We choose to use a 30%, 60% and 90% probability intervals.
[40] The EU target for 2020 is below 10% for early school leavers and at least 40% for tertiary education attainment. For the national targets, see http://ec.europa.eu/europe2020/pdf/targets_en.pdf.

The results of the current forecasting round are illustrated using a qualitative ranking scale in order not to overestimate the implications of our approach. Having worked with such a limited data set, the best practice was to avoid giving point forecasts and instead to highlight the uncertainty as a separate outcome of this type of exercise. This approach follows the discussion in Tay and Wallis (2000) concerning the production and presentation of conditional density forecasts.

All EU27 countries were assessed according to the probability of reaching the targets based on the forecasted 2020 distribution probability. More detailed country specific forecasts for both benchmarks, including fan-charts for a better illustration of the inherent uncertainty, can be found in Annex 1 and Annex 2 at the end of this paper.

According to our classification:

- *high* probability means that 80% of the expected outcomes would lie below the target in the case of *early school leavers* / above the target in case of *tertiary education attainment*.
- *quite high* probability corresponds to between 65% and 80% of the expected outcomes.
- *fair* probability corresponds to between 35% and 65%. This means that the (national or EU) target would be included in this interval, but we still cannot be too confident whether the outcome would be really below target in the case of *early school leavers* / above target in the case of *tertiary education attainment*.
- *quite low* probability corresponds to between 20% and 35% of the expected outcomes.
- *low* probability corresponds to less that 20% of the expected outcomes lying below target in the case of *early school leavers* / above target in the case of *tertiary education attainment*.

An overview of the forecasts' results could be observed in Table 5 below, where ELE is the abbreviation for *early school leaving* and TEA – for *tertiary education attainment*.

Table 5

	2012 data		National target		2020 probability of ELE reaching …		2020 probability of TEA reaching …	
	ELE	TEA	ELE	TEA	National target	EU target	National target	EU target
AT	7.6	26.3	9.5	38(*)	High	High	Quite low(*)	Low
BE	12	43.9	9.5	47	Fair	Fair	Fair	High
BG	12.5	26.9	11	36	Quite high	Fair	Fair	Quite low
CY	11.4	49.9	10	46	Quite high	Quite high	Quite high	High
CZ	5.5	25.6	5.5	32	Fair	High	Fair	Quite low
DE	10.5	31.9	10	42(*)	Fair	Fair	Quite low(*)	Quite low
DK	9.1	43	10	40	Quite high	Quite high	Quite high	Quite high
EE	10.5	39.1	9.5	40	Fair	Fair	High	High
ES	24.9	40.1	15	44	Quite low	Low	Quite high	Quite high
FI	8.9	45.8	8	42(**)	Fair	Quite high	High	High
FR	11.6	43.6	9.5	50(^)	Quite high	Quite high	Fair	Quite high
EL	11.4	30.9	9.7	32	Quite high	Quite high	Quite high	Fair
HU	11.5	29.9	10	30.3	Fair	Fair	Quite high	Fair
IE	9.7	51.1	8	60	Quite high	High	Fair	High
IT	17.6	21.7	16	26	Quite high	Low	Fair	Low
LT	6.5	48.7	9	40	Quite high	High	High	High
LU	8.1	49.6	10	40	High	High	High	High
LV	10.5	37	13.4	34	Quite high	Fair	High	High

MT	22.6	22.4	29(+)	33	High	Low	Quite low	low
NL	8.8	42.3	8	40	Quite high	High	Quite high	Quite high
PL	5.7	39.1	4.5	45	Fair	High	Quite high	High
PT	20.8	27.2	10	40	Quite low	Quite low	Quite low	Quite low
RO	17.4	21.8	11.3	26.7	Quite low	Low	Quite high	Quite low
SE	7.5	47.9	10	40	High	High	High	High
SI	4.4	39.2	5	40	High	High	Quite high	Quite high
SK	5.3	23.7	6	40	Quite high	High	Quite low	Quite low
UK	13.5	47.1	-	-	-	Fair	-	High

(*) For AT and DE, the national targets include postsecondary attainment (ISCED 4/4A for AT and ISCED 4 for DE), which means that our forecasts are underestimating the countries' ability to reach their national targets in terms of tertiary education attainment.

(**) For FI the national target is defined more narrowly that the EU target and excludes technological institutes.

(^) For FR the national target for TEA refers to the age group 17-33 years old.

(+) For MT the national target for ELE was based on data previous to the 2012 revision.

Note: In case that national targets were set as intervals, they were approximated up to the most conservative value.

The forecasts and the associated probability evaluation depicted in Table 5 must be interpreted with due care and under the assumptions presented above. Obviously, one can add his or her own country-specific inputs into the evaluation, given better knowledge about the expected policy reforms or significant drivers which have been omitted in our selected empirical specifications. For example migration flows were not accounted for in our empirical models, but were highlighted when extrapolating the adults' education attainment up to 2020 (see Annex 5, footnote 46) as a potential driver of the forecast in an alternative scenario. We also assume no role for latter training or participation in the education for parental cohorts; relaxing this assumption would also mean that policies encouraging lifelong learning could make a difference and alter our current evaluation.

In general, Table 5 paints a quite optimistic picture for the majority of the Member States regarding the 2020 probability of reaching national and EU targets in education. According to our assessment 16 Member States have a high or quite high probability to reach their national targets for early school leavers, while 7 other Member States have a fair probability. Evaluating Member States against EU target provides more or less the same distribution of probabilities. In the case of tertiary education attainment 15 Member States have a high or quite high probability of reaching their national targets by 2020, while 6 have a fair probability. We prefer to exclude Austria and Germany from this assessment as long as national targets include postsecondary attainment, so that our evaluation is certainly an underestimation. The national target for tertiary education is differently defined in Finland and France, while Malta is about to revise it. As some Member States have set less ambitious national targets compared with the EU target, the evaluation is certainly less optimistic in this case, as shown in the last column of Table 5.

5. A birth cohort model for tertiary education attainment

This section provides a cross-check of the results on tertiary educational attainment projected by the econometric model and detailed in the preceding section. The concept of cohorts is widely employed in the research on demographics, fertility, health and education to analyse a targeted population with certain characteristics born in a specific period (see Yang and Land, 2013; Myrskylä et al., 2013; McGarry et al. 2012; and Ni Bhrolchain, 1992). In the current report, we use birth cohorts to assess the education level pertaining to individuals, who will be in the target population of the benchmark, i.e. 30-34 years old.

The specific birth cohort methodology, which is employed to calculate the benchmark value for the coming years, has been developed by CRELL and constitutes on-going research[41]. While the forecasting method relies on an econometric model to predict future benchmark values, the birth cohort model uses administrative data from the UIS/OECD/EUROSTAT data collection (UOE) to construct an indicator measuring the completion rates of the target population. This is done by looking at new entrants to tertiary education by age group, the average duration of studies and the average completion rate as reported by countries. Using these variables we are able to track individuals in different birth cohorts and based on the flow of new entrants in tertiary education, we calculate the output, i.e. the proportion, in the coming years, of people aged between 30 and 34 with completed tertiary education. There are seven steps involved in the calculation of the headline target, which are discussed below. Data for the calculation of the birth cohort method is mainly drawn from UOE and OECD (see Table 6).

Step 1: Identifying the cohorts counted into the headline indicator
For each year of the headline indicator to be calculated, we have identified the respective birth cohorts (all individuals born in a certain year, for example in 1990) entering into the indicator.

Table 6: Data sources and variables, cohort approach

Variable	Data Source	Definition	Availability
Number of individuals that entered tertiary education by age group (ISCED 5A and B)	UOE	New entrants to the tertiary education type 5A and 5B are students who are entering any program leading to a recognized qualification at the 5A or 5B level of education for the first time, irrespective of whether the students enter the program at the beginning or at an advanced stage of the program. New entrants are between 18 and 29 years old.	1998-2011, but missing values for several countries
Population by age group	UOE	Total population by age group for the ages 18 – 29	1998-2011
Completion rate	Eurostat, 2009 and OECD, Education at a Glance, 2010	Completion rates in tertiary-education type 5A represent the proportion of those who enter a tertiary-education type 5A programme and go on to graduate from at least a tertiary-education type 5A program.	2008
Average duration of studies	OECD Indicators, Education at a Glance, 2008	Average duration of tertiary studies (in years) for ISCED 5A and 5B.	2008

Table 7 below shows in the last line 5 different birth cohorts, namely those born between 1986 and 1990. Those are the birth cohorts contributing to the calculation of the 2020 target value. For example, a person born in 1986 will be 34 years old in 2020 and hence will be counted in the benchmark (1986+34years=2020). Likewise a person born in 1990 will be 30 years old in 2020 and thus will be counted as well. Since one has to monitor the headline target starting from 2013 until 2020, the birth cohorts considered are those between 1979 and 1990.

Table 7: Birth cohorts entering in the calculation of each year of the target

Target value	Birth cohorts:
2011	1977, 1978, 1979, 1980, 1981
2012	1978, 1979, 1980, 1981, 1982
2013	1979, 1980, 1981, 1982, 1983
2014	1980, 1981, 1982, 1983, 1984
…	…
2019	1985, 1986, 1987, 1988, 1989
2020	1986, 1987, 1988, 1989, 1990

[41] See Badescu et al. (2012).

Step 2: Extrapolation of missing data
As for all data analysis, the validity of the birth cohort model and thereby the calculated benchmark values depend critically on having the most complete information on students entering tertiary education. Standard extrapolation methods to impute missing data could add value to our approach. However, due to the structure of the data on new entrants, most of the missing values would affect the forecast mostly between 2011-2013 and 2017-2020. In Annex 2 we only present those values for which we have most complete data and which we consider as most reliable, leaving aside here any issues related to the imputation procedure[42] derived uncertainty.

Step 3: Computing the number of new entrants to tertiary education by birth cohort
For each birth cohort, we calculate the number of individuals that entered tertiary education, i.e. ISCED 5A and 5B. In order to take into account that some students of one birth cohort might decide to enter into tertiary education when they are older than 18, we considered all entries between 18 and 29. Beyond 29 the entries are negligible (less than 1%) and thus not included. Table 7 shows the birth cohorts included in the simulation.

Summing up, Steps 1 to 3 allow us to calculate, for each year of the target, the total number of individuals aged 30 to 34 that enrolled in tertiary education. For the 2020 target value, it corresponds to the total number of individuals born between 1986 and 1990 who participated in tertiary education when they were aged between 18 and 29 years old.

Step 4: Computing the total population aged 30 to 34 for each year towards the target
For each year up to 2020, we calculate the number of individuals belonging to the cohorts that are counted in the benchmark. For example, for the benchmark value 2020, we compute the total population of individuals born between 1977 and 1990. Annex 6 explains the calculations in greater detail.

Step 5: Calculating the ratio of new entrants to birth cohort population
By dividing the number of corrected new entrants of each cohort by the total birth cohort population, we obtain a cohort-specific gross entry rate in tertiary education (GER henceforth).

Step 6: Computing the mean of the gross entry rate for each target year
From Step 4, we obtain the GERs for five different cohorts per target year. The overall gross entry rate is then calculated by taking the mean of the five cohort-specific GERs[43].

Step 7: Computing the target value
Last, we multiply the mean GER by the country-specific completion rate. Unfortunately, we do not have a completion rate for each cohort. The completion rate used for the forecast is the completion rate for the tertiary education level 5A available from Eurostat for the year 2008 (see information given in Table 6). This implies that the calculated values of the benchmark crucially depend on having a constant tertiary completion rate across years and types of programs (5A and 5B)[44].

The calculated benchmark values for the most reliable years, i.e. 2014-2017, are depicted together with the country-specific forecasts in Annex 2. As can be seen from the charts in Annex 2, the calculated benchmark values from the birth cohort are located *within the*

[42] On-going research at JRC/CRELL aims at quantifying how much of the variations in the forecasted values of the benchmark are due to the imputation procedure and on conducting an uncertainty analysis which allows providing confidence bounds of the model output.

[43] Alternatively, we could use the weighted mean GER, with the weight being given by the proportion of the population belonging to each cohort.

[44] To provide some validation of the birth cohort model, we compare our calculated benchmark value for 2011 with the 'true' benchmark value as provided by Eurostat. Due to limitations of the data on new entrants from the UOE, the earliest year for which we can calculate the benchmark value is 2011.

confidence bounds of the forecast. This provides strong evidence for the validity of the approach.

Note that differences in the exact location of the birth cohort values *vis-a-vis* the forecast bounds might stem from the specific assumptions underlying the birth cohort model. In particular, the birth cohort model assumes a constant completion rate and invariant average duration of studies. These assumptions might hence lead to a more conservative or more optimistic calculated benchmark value compared to the forecast derived from the empirical model.

6. Conclusions

This paper proposes an econometric methodology for forecasting the twofold Europe 2020 headline target using a cross country time series approach. We take the human capital theory as a starting point in our analysis in order to understand what drives *early school leaving* and *tertiary education attainment* dynamics over time. We rely on the empirical literature to identify the best proxies and estimate the models in a panel setting to better deal with a limited dataset.

We construct country-specific forecasts under simple but realistic assumptions about the expected level of adults' educational attainment and given the other underlying determinants of schooling decisions uncovered by our empirical analysis. In order to enable comparability and policy analysis, a uniform methodology was applied to all Member States. 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.

Our forecasts paint a favourable outlook across the European Union in terms of reaching the targets set for the two benchmarks. More than half of the EU27 Member States show a high or quite high probability of reaching the targets by 2020 for both benchmarks. About a quarter of the Member States have forecasts which are very close to their specific target, meaning that the outcome could be either above or below the target value. For the remaining countries, our exercise shows an under-performance in terms of reaching the targets, which is likely to require more in-depth investigation into the determinants of tertiary education attainment or early school leavers, depending on the country.

The forecasts tell us how early school leaving and tertiary education attainment are likely to develop over the next years *if nothing changes in terms of policy measures*. The strong assumption of *no policy change* suggests that implementation of adequate reforms could therefore lead to a revision of our evaluation.

Our empirical analysis highlights some key areas where policy impact is likely to be larger or more efficient. Family background is an important determinant and could therefore be a useful lever for policy action. Some of the options that might be considered in this sense refer to encouraging lifelong learning, removing barriers to participate in training programs or introducing outreach programmes towards the disadvantaged.

Apart from family influences, our benchmarks are driven by schooling decisions of young individuals. For early school leavers, adequate counselling about long term benefits of education and related job prospects are likely to help alleviate the risks emanating from business cycle fluctuations[45]. Tertiary education attainment is likely to benefit from economic incentives arising from a more flexible and efficient labour market. Reducing skill mismatches and improving the transmission of productivity gains into real wages via structural reforms could reduce the uncertainty around future *skill premium*, clearing the way for better schooling decisions.

[45] High entry-level wages or labour shortages characterized by a high labour demand for workers with low skills or lacking experience could have undesired effects on education decisions.

Our analysis has some inherent limitations, mainly related to data availability, which hinders a full investigation of the determinants of education decisions. Yet, the added value of this approach rests in its ability to integrate judgements, analysis and alternative scenarios from country experts, such as those addressing omitted influences or anticipated policy changes. In addition, from a methodological point of view, our approach could be extended in the future by aggregating the information provided by various empirical specifications, such as in a model-averaging strategy.

References

Acemoglu, D. 1996. A Microfoundation for Social Increasing Returns in Human Capital Accumulation. *The Quarterly Journal of Economics* , Vol. 111, No. 3 (Aug., 1996), pp. 779-804

Acemoglu, D. 2000. Technical Change, Inequality, and the Labor Market. NBER Working Paper No. 7800

Acemoglu, D. and S. Johnson, 2007. Disease and Development: The Effect of Life Expectancy on Economic Growth, Journal of Political Economy, University of Chicago Press, vol. 115(6), pages 925-985, December.

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

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

Badescu, M., B. d'Hombres and A. Weber, 2012. Monitoring the EU headline targets in education and training. Methodology and country sheets, JRC Scientific and Technical Reports

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

Ben Porath, Y. 1967. The Production of Human Capital and the Life Cycle of Earnings, JP E 75 (1967), pp. 352-365.

Becker, G. 1964, Human Capital, New York: Columbia University Press for the National Bureau of Economic Research.

Becker, G.S. 1975. Human Capital, 2nd edition, New York, Columbia University Press

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

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

Betts, J., and L. McFarland 1995. Safe port in a storm. The impact of labor market conditions on community college enrollments. Journal of Human Resources, 30 (4), 741–765.

Bils M. and P. J. Klenow 2000. Does Schooling Cause Growth?, American Economic Review, American Economic Association, vol. 90(5), pages 1160-1183, December

Bjørnstad, R. and M. L. Gjelsvik 2011, How are Educational Groups Affected by Economic Shocks and Trends?, Reports 50/2011, Statistics Norway

Buchinsky, M. and P. Leslie. 2010. Educational Attainment and the Changing U.S. Wage Structure: Dynamic Implications on Young Individuals' Choices. Journal of Labor Economics Vol. 28, No. 3 (July), pp. 541-594

Bureau of Labor Statistics, U.S. Department of Labor, *Occupational Outlook Handbook, 2012-13 Edition*, Projections Overview, on the Internet at http://www.bls.gov/ooh/about/projections-overview.htm

Cameron, S.V. and J.J. Heckman. 2001. The Dynamics of Educational Attainment for Black, Hispanic, and White Males, Journal of Political Economy, 2001, vol. 109, no. 3

Campos, N. F., G. Hughes, S. Jurajda and D. Munich 1999. When the Future is not What it Used to Be: Lessons from the Western European Experience to Forecasting Education and Training in Transition Economies. William Davidson Institute (WDI) - Working Papers 265

Drukker, D. M. 2003. Testing for Serial Correlation in Linear Panel-Data Models, Stata Journal 3: 168–177

Eggert, W., T. Krieger and V. Meier 2010. Education, Unemployment and Migration. Journal of Public Economics 94 354–362

Ermisch, J. and M. Francesconi, 2000. Family Matters: Impacts of Family Background on Educational Attainments. Economica 68, 137-156

European Commission 2010, Efficiency and Effectiveness of Public Expenditure on Tertiary Education in the EU: Joint Report by the Economic Policy Committee and Directorate-General for Economic and Financial Affairs available at http://ec.europa.eu/economy_finance/publications/occasional_paper/2010/pdf/ocp70_en.pdf

Grubb, W. 1988. Vocationalizing Higher Education: the causes of enrollment and completion in public two-year colleges, 1970–1980. Economics of Education Review, 7 (3), 301–319.

Hazan, M. 2009. Longevity and Lifetime Labor Supply: Evidence and Implications. Econometrica 77, 1829-1863.

Haveman R. and B. Wolfe 1995. The Determinants of Children's Attainment: A Review of Methods and Findings, Journal of Economic Literature 32, 1829-78

Heckman, J.J. 1976. A Life-Cycle Model of Earnings, Learning, and Consumption. Journal of Political Economy Vol. 84, No. 4, The University of Chicago Press

Heckman, J.J., L. Lochner, C. Taber 1998. Human Capital Formation and General Equilibrium Treatment Effects: A Study of Tax and Tuition Policy. American Economic Review, Vol. 88, no. 2, 381-386

Jacoby H.G. 1994. Borrowing Constraints and Progress Through School: Evidence from Peru. The Review of Economics and Statistics. Vol. 76, No. 1 (Feb., 1994), pp. 151-160

Katz L. F. and D. H. Autor 1999. Changes in the wage structure and earnings inequality. *Handbook of labor economics*, *3*, 1463-1555.

Keane, M.P. and K.I. Wolpin. 2001. The Effects of Parental Transfers and Borrowing Constraints on Education Attainment. International Economic Review, Vol 42, No. 4 (November)

Kodde, D. 1986. Uncertainty and the Demand for Education, Review of Economics and Statistics, 68 (3), 460-467

Lutz, W. and S. Scherbov, 2006. Global Age-specific Literacy Projections Model (GALP): Rationale, Methodology and Software, UNESCO Institute for Statistics, Montréal (Québec), Canada

Mattila, J. P. 1982. Determinants of Male School Enrollment: A Time Series Analysis. Review of Economics and Statistics, 64 (2), 242-251

Mincer, J. 1974, Schooling Experience and Earnings NY: NBER Columbia U Press.

Micklewright, J., M., Pearson and S. Smith 1990 - The Economic Journal , Vol. 100, No. 400, Conference Papers(1990), pp. 163-169

Mikko Myrskylä, M, J. R. Goldstein, and Y.-H. A. Cheng, 2013: New Cohort Fertility Forecasts for the Developed World, Revised version 2013, MPIDR WORKING PAPER WP 2012-014

McGarry, L., V.S. Pawar, H.R. Panchmatia, J.L. Rubin, G.L. Davis, Z.M. Younossi, J.C. Capretta, M.J. O'Grady, and M.C. Weintsein, 2012: Economic model of a birth cohort screening program for hepatitis C virus, Hepatology 55 (5),1344-55

Mocetti, S. 2010. Educational Choices and the Selection Process: Before and After Compulsory Schooling. Education Economics, Vol. 20, No. 2, 189-209

Murphy K.M. and F. Welch 1992. The Structure of Wages. The Quarterly Journal of Economics. Vol. 107. No 1, 285-326

Ni Bhrolchain, M., 1992: Period Paramount? A Critique of the Cohort Approach to Fertility, Population and Development Review, 18 (4), 599-629

Pissarides C. 1981. Staying-on at school in England and Wales. Economica, 48 (192), 345-363

Petrongolo, B. and M. J. S. Segundo, 2002. Staying-on at school at 16: the impact of labor market conditions in Spain. Economics of Education Review, Volume 21, Issue 4, Pages 353-365

Restuccia, D. and G. Vandenbroucke. 2013. Explaining Education Attainment across Countries and over Time. Working Paper 469, University of Toronto, Department of Ecponomics

Rice, P.G. 1987. The Demand for Post Compulsory Education in the UK and the effects on educational maintenance allowances. Economica 54, 465-75

Roodman, D. 2006. How to do xtabond2: an introduction to "Difference" and "System" GMM in Stata. Center for Global Development Working Paper Number 103

Tay, A.S. and K.F. Wallis, 2000. Density forecasting: a survey. Journal of Forecasting, 19, 235-254

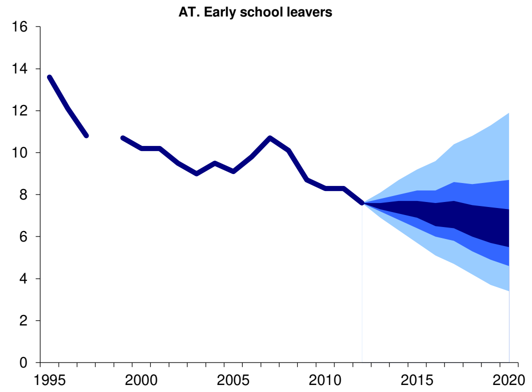
Tieben, N. and M. Wolbers, 2010. Success and failure in secondary education: socio-economic background effects on secondary school outcome in the Netherlands, 1927–1998. British Journal of Sociology of Education, 31:3, 277-290

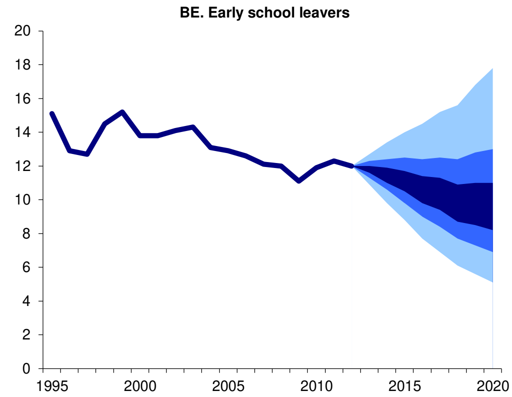
Whitfield, K. and R. A. Wilson 1991. Staying-on in Full Time Education: The Education Participation Rate of 16-years-olds. Economica, New Series, Vol. 58, No. 231 August, pp. 391-404

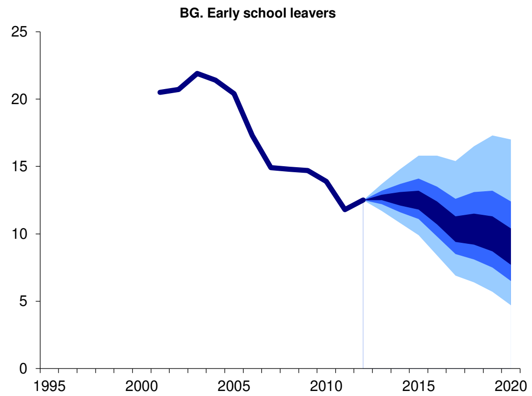
Wooldridge, J. M. 2002, *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

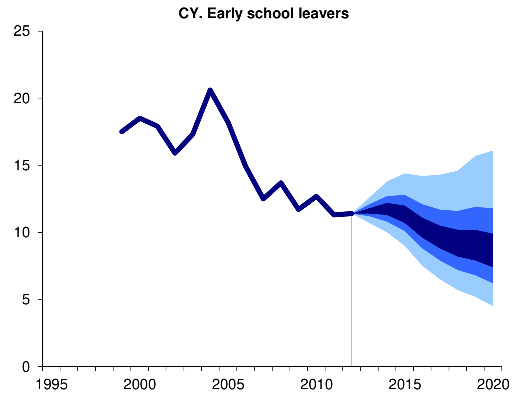
Yang, D.Y, and K. C. Land, 2013: Age-Period-Cohort Analysis: New Models, Methods, and Empirical Applications, Chapman and Hall/CRC; Taylor & Francis Group, Boca Raton

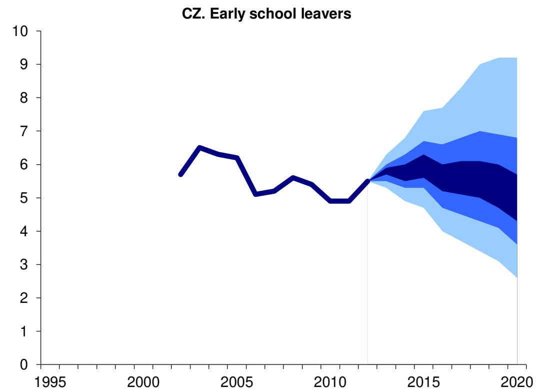
Annex 1. Country specific forecasts and uncertainty intervals for *early school leavers*

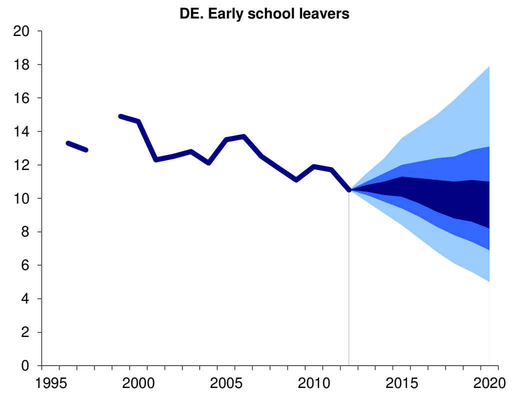




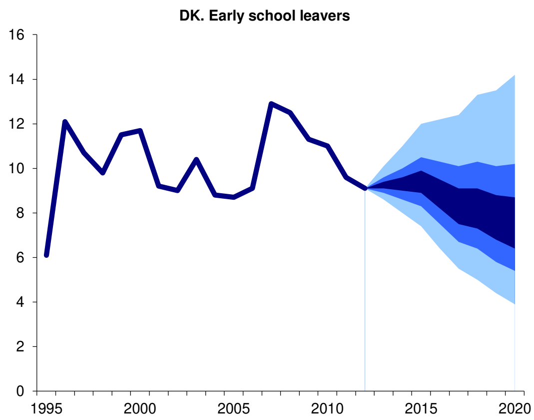


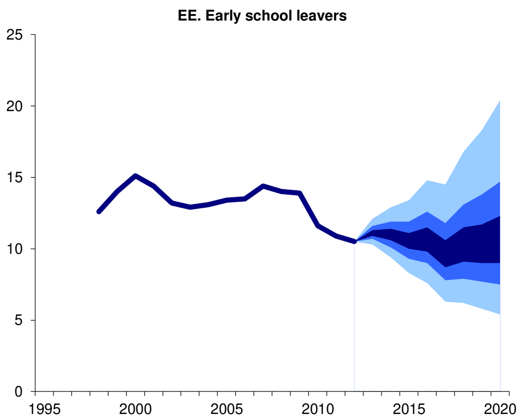


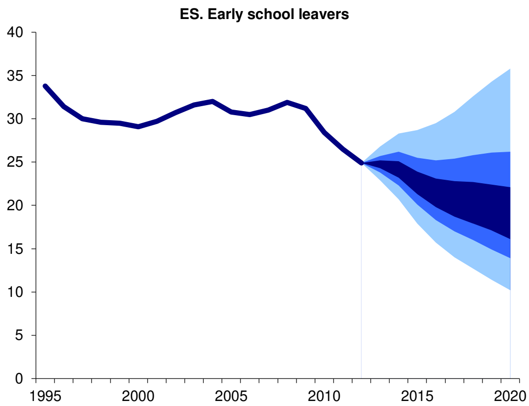


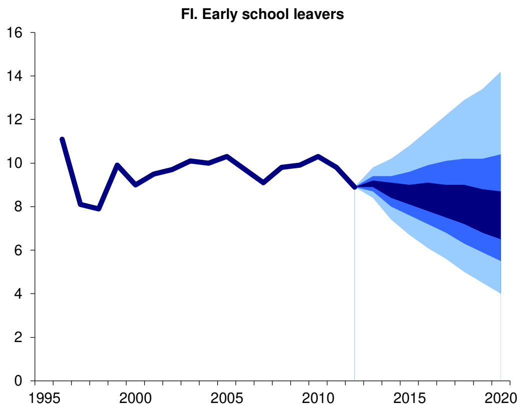


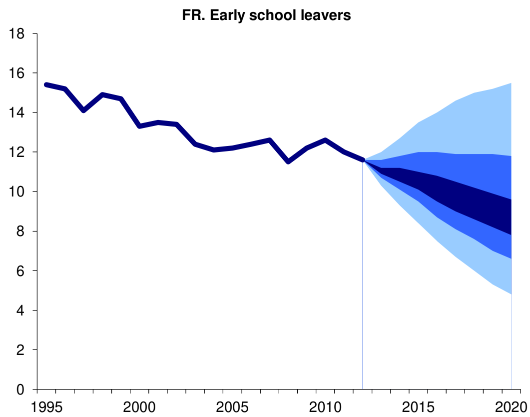
Note: The shaded areas in the figures above plot the 30% (dark blue), 60% (blue) and 90% (light blue) confidence intervals for *early school leavers* forecast constructed using the best econometric specification presented in section 4.1. More certainty would mean wider forecast range. Based on these confidence intervals we have constructed the qualitative forecasts presented in section 5.
Please note that the scale of the vertical axis might be different depending on the country.

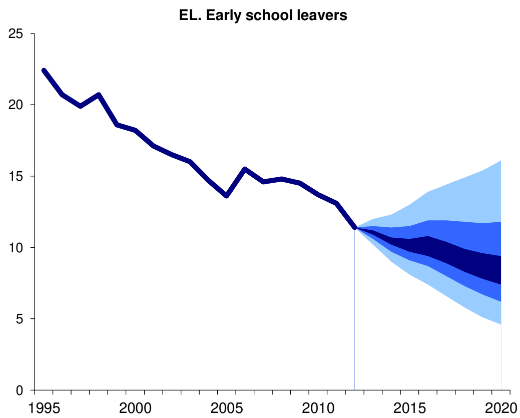




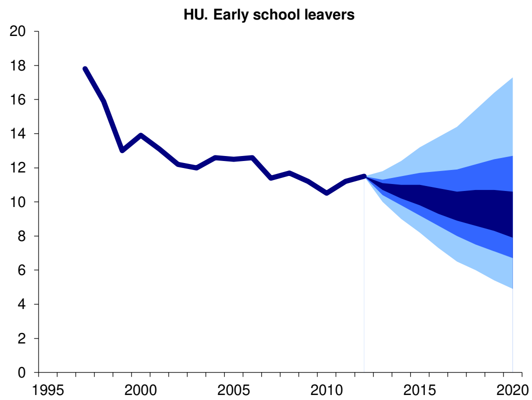


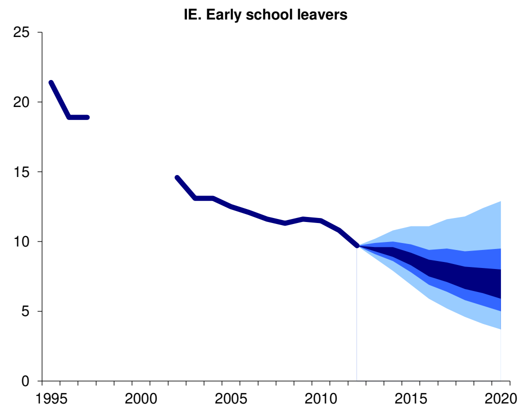


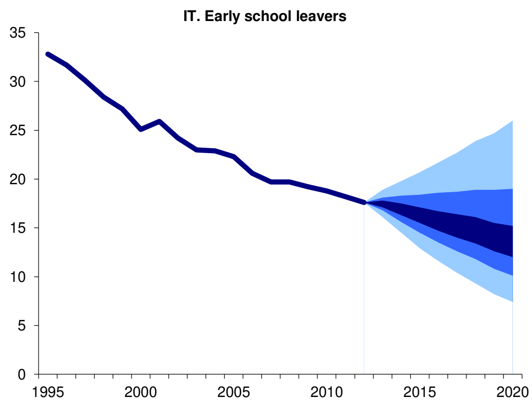


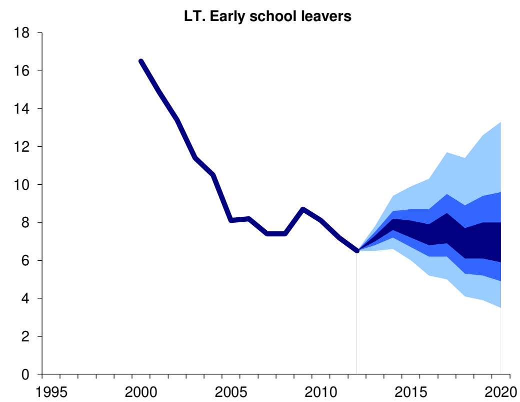


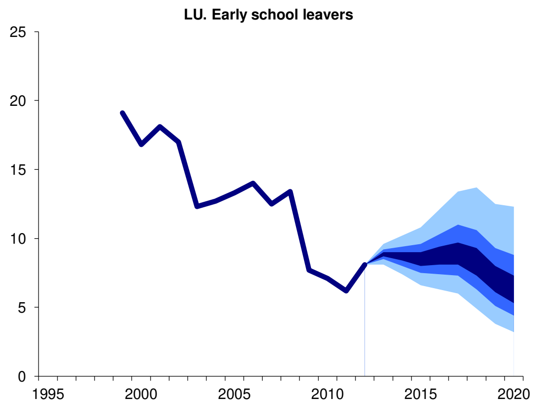
Note: The shaded areas in the figures above plot the 30% (dark blue), 60% (blue) and 90% (light blue) confidence intervals for *early school leavers* forecast constructed using the best econometric specification presented in section 4.1. More certainty would mean wider forecast range. Based on these confidence intervals we have constructed the qualitative forecasts presented in section 5.
Please note that the scale of the vertical axis might be different depending on the country.

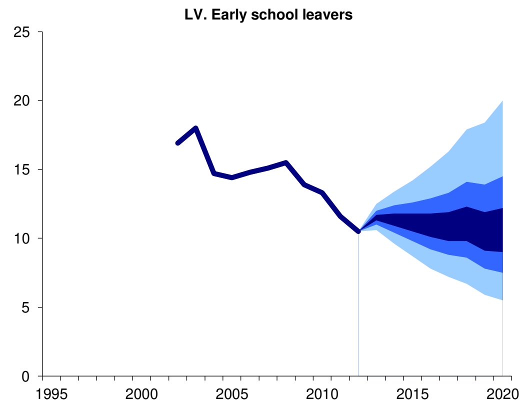






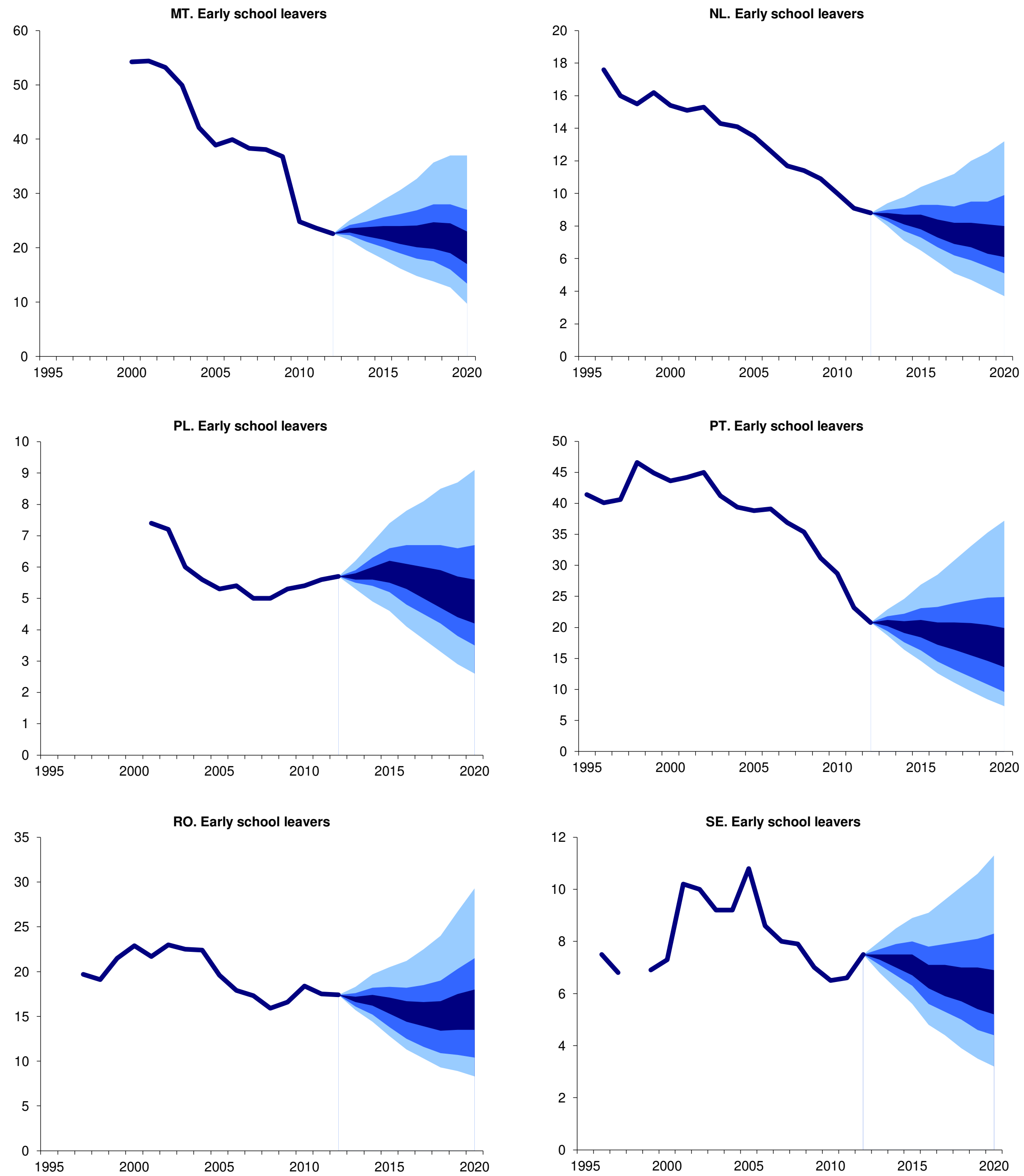






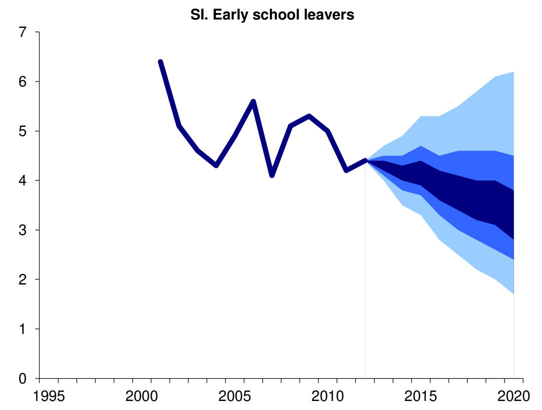
Note: The shaded areas in the figures above plot the 30% (dark blue), 60% (blue) and 90% (light blue) confidence intervals for *early school leavers* forecast constructed using the best econometric specification presented in section 4.1. More certainty would mean wider forecast range. Based on these confidence intervals we have constructed the qualitative forecasts presented in section 5.

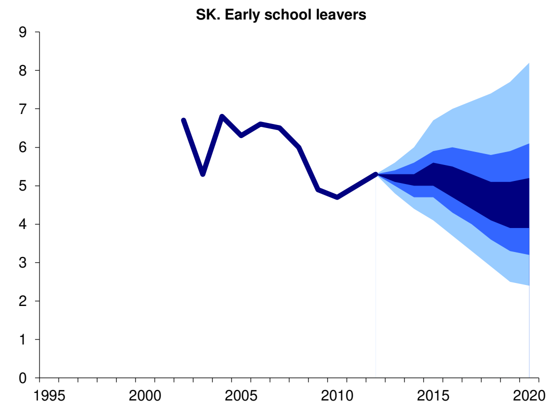
Please note that the scale of the vertical axis might be different depending on the country.

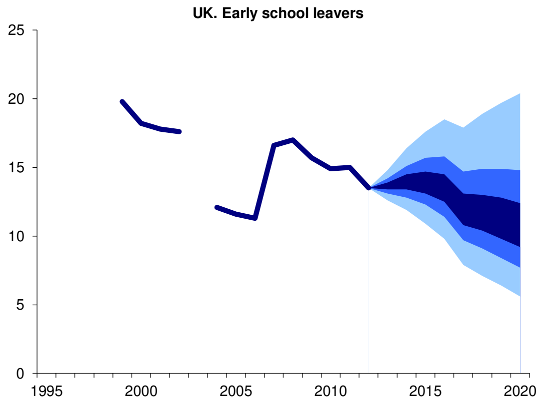


Note: The shaded areas in the figures above plot the 30% (dark blue), 60% (blue) and 90% (light blue) confidence intervals for *early school leavers* forecast constructed using the best econometric specification presented in section 4.1. More certainty would mean wider forecast range. Based on these confidence intervals we have constructed the qualitative forecasts presented in section 5.

Please note that the scale of the vertical axis might be different depending on the country.

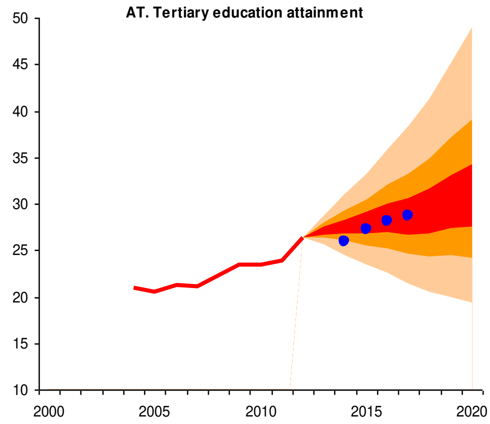


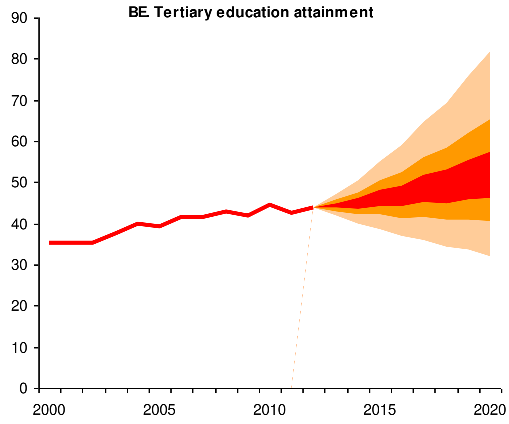


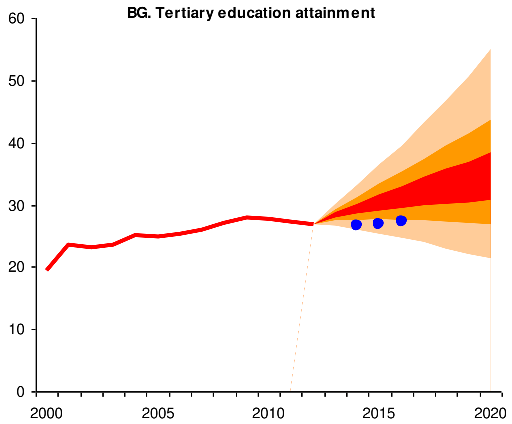


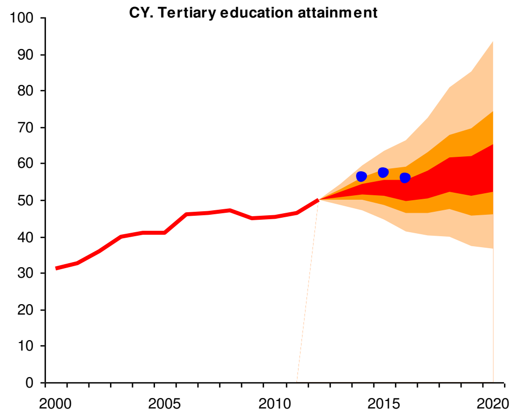
Note: The shaded areas in the figures above plot the 30% (dark blue), 60% (blue) and 90% (light blue) confidence intervals for *early school leavers* forecast constructed using the best econometric specification presented in section 4.1. More certainty would mean wider forecast range. Based on these confidence intervals we have constructed the qualitative forecasts presented in section 5.
Please note that the scale of the vertical axis might be different depending on the country.

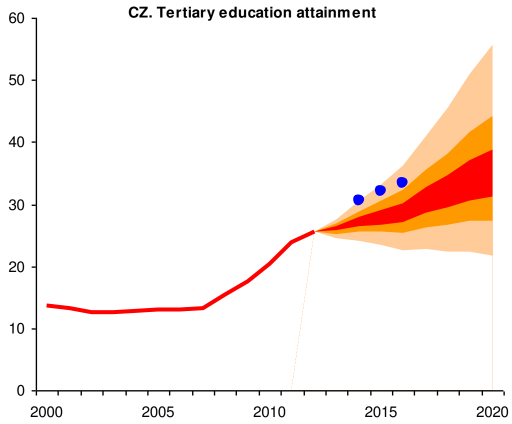
Annex 2. Country specific forecasts and uncertainty intervals for *tertiary education attainment*

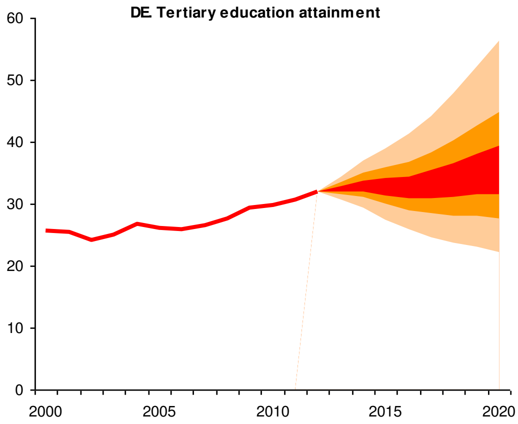




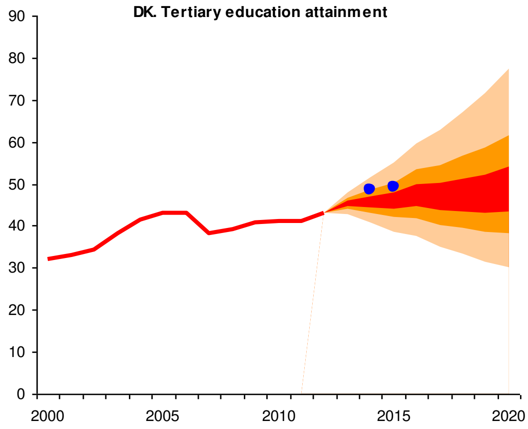


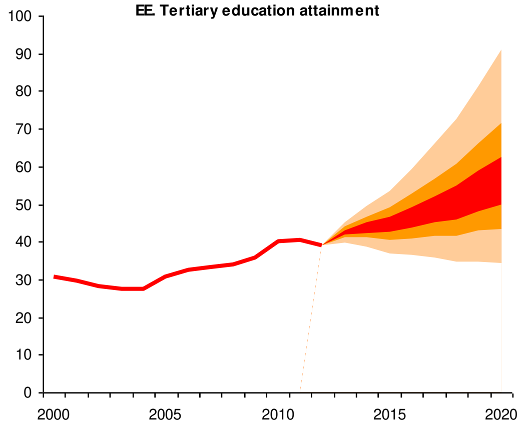


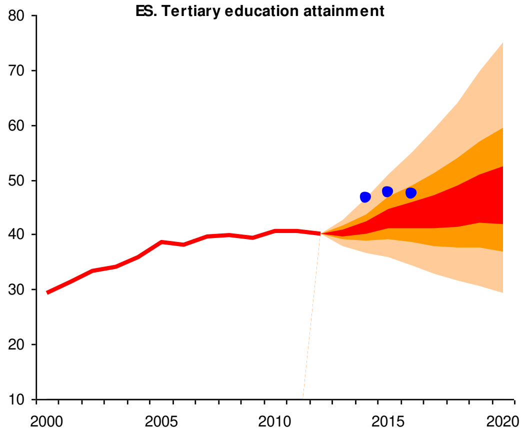


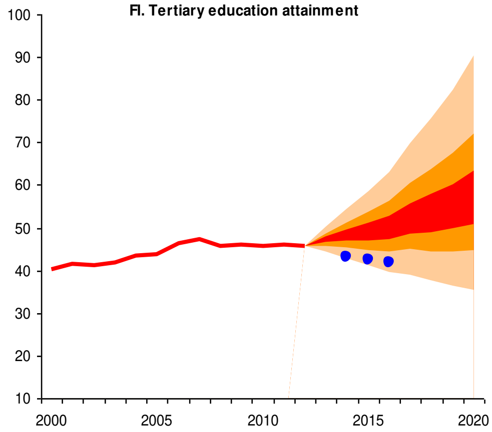


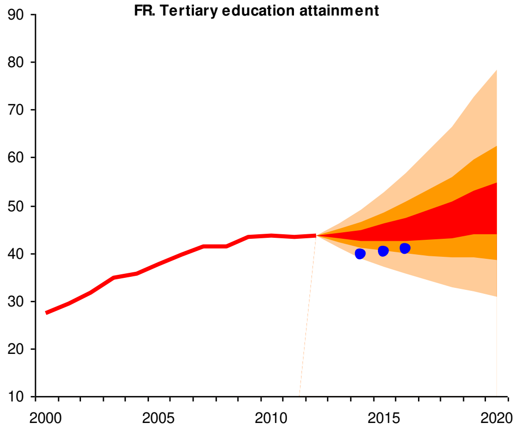
Note: The shaded areas in the figures above plot the 30% (dark red), 60% (red) and 90% (light red) confidence intervals for *tertiary education attainment* forecast constructed using the best econometric specification presented in section 4.1. More certainty would mean wider forecast range. Based on these confidence intervals we have constructed the qualitative forecasts presented in section 5. The blue points represent the forecasts obtained from the birth-cohort model described in section 6.
Please note that the scale of the vertical axis might be different depending on the country.

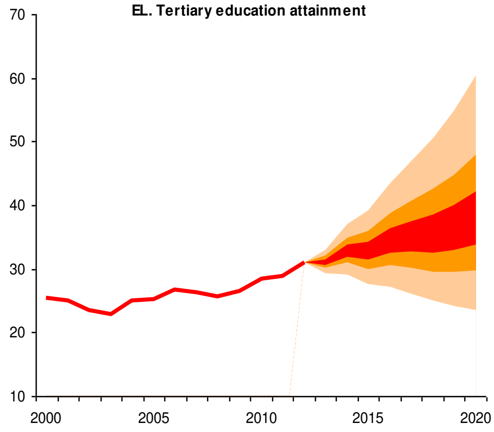




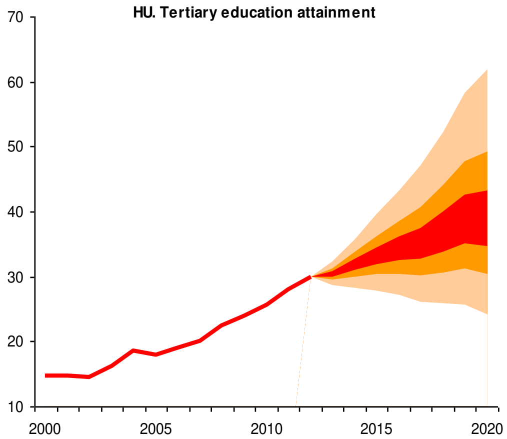


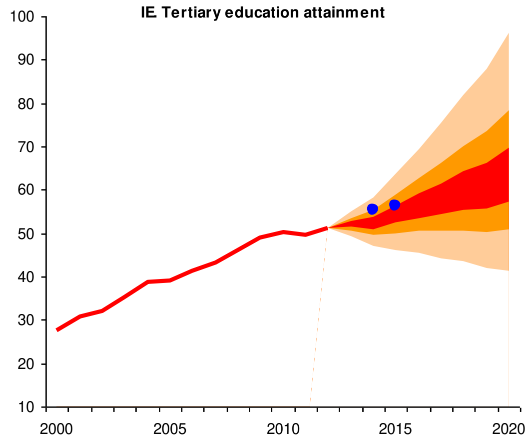


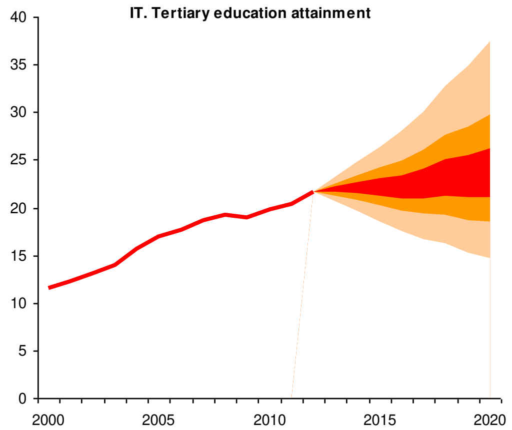


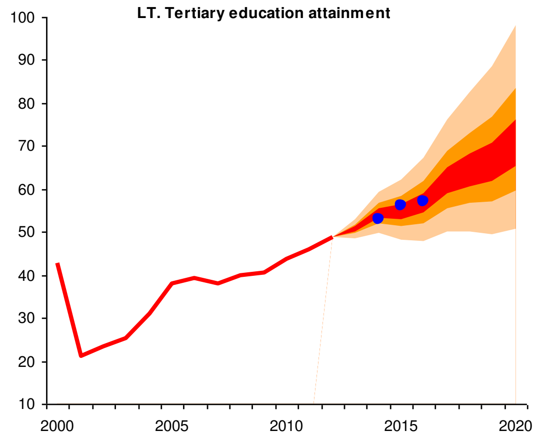


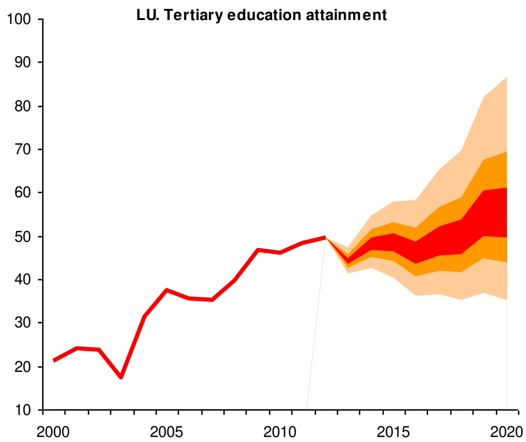
Note: The shaded areas in the figures above plot the 30% (dark red), 60% (red) and 90% (light red) confidence intervals for *tertiary education attainment* forecast constructed using the best econometric specification presented in section 4.1. More certainty would mean wider forecast range. Based on these confidence intervals we have constructed the qualitative forecasts presented in section 5. The blue points represent the forecasts obtained from the birth-cohort model described in section 6.
Please note that the scale of the vertical axis might be different depending on the country.

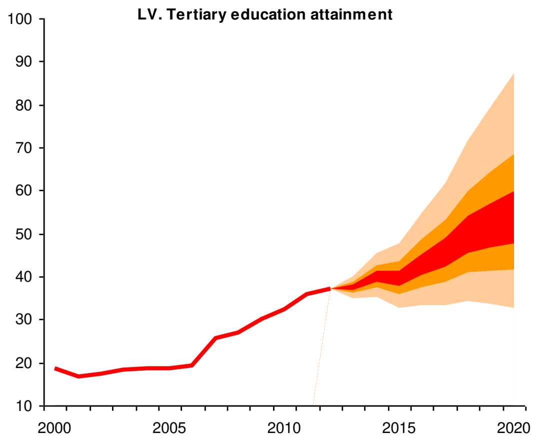




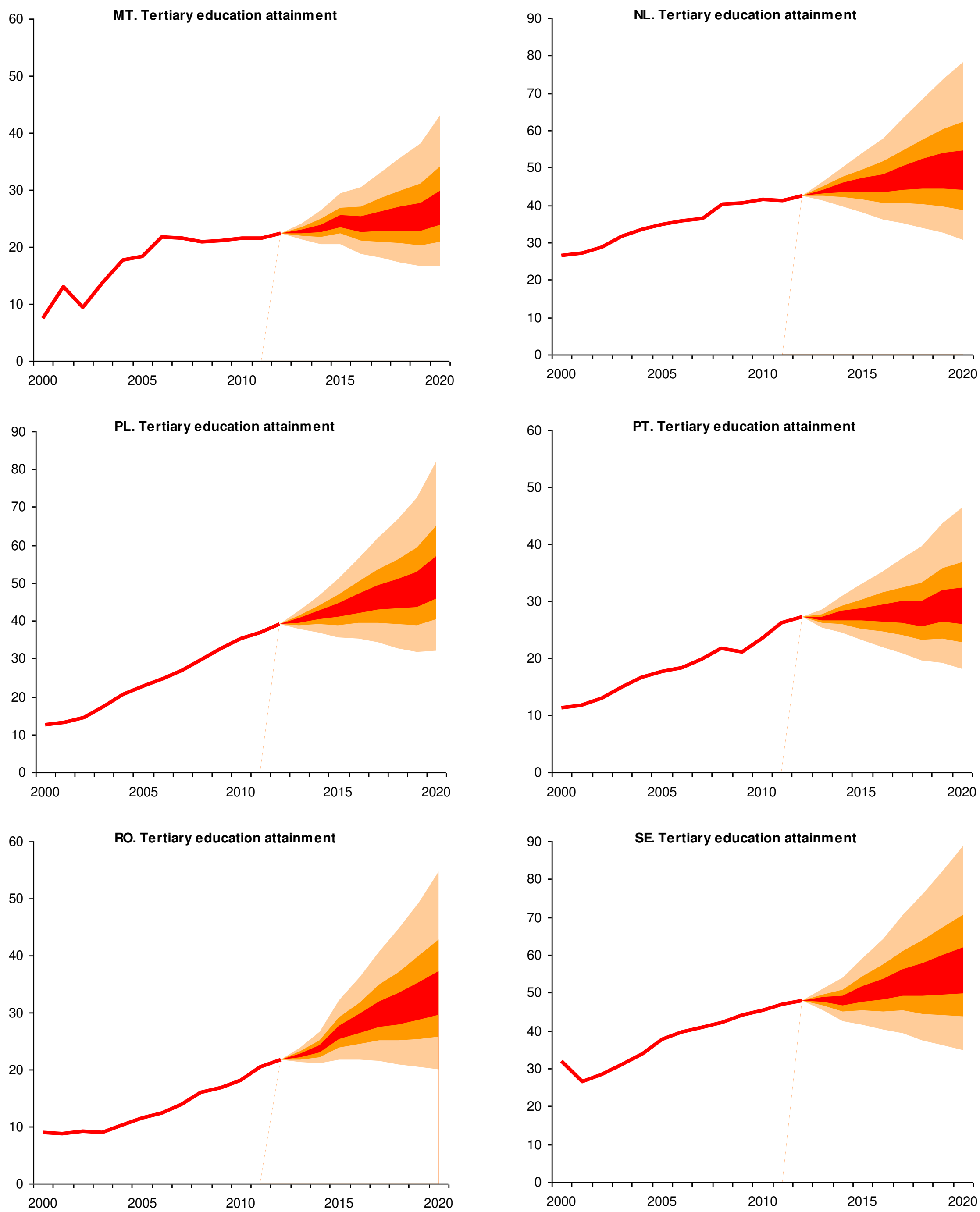






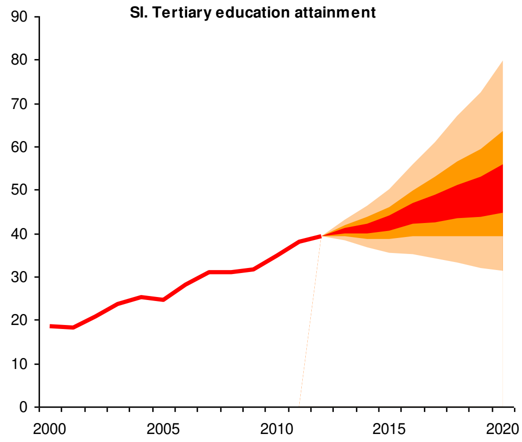


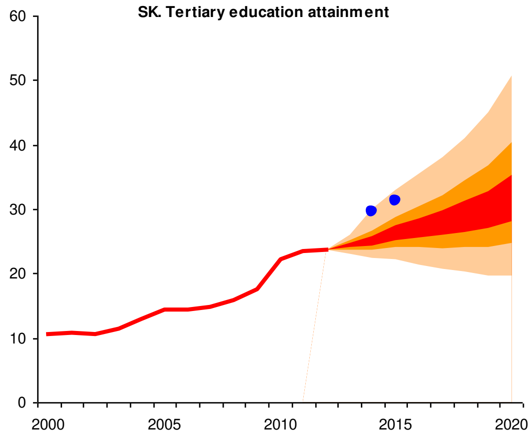
Note: The shaded areas in the figures above plot the 30% (dark red), 60% (red) and 90% (light red) confidence intervals for *tertiary education attainment* forecast constructed using the best econometric specification presented in section 4.1. More certainty would mean wider forecast range. Based on these confidence intervals we have constructed the qualitative forecasts presented in section 5. The blue points represent the forecasts obtained from the birth-cohort model described in section 6.
Please note that the scale of the vertical axis might be different depending on the country.

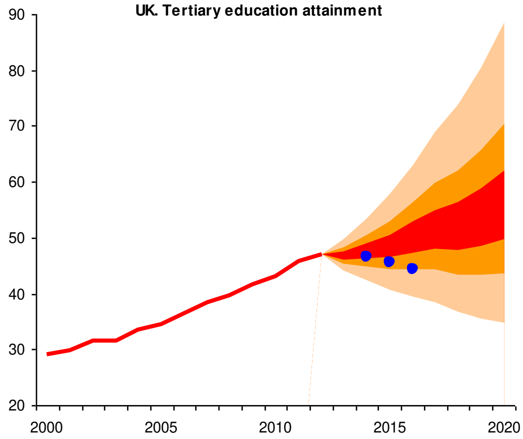


Note: The shaded areas in the figures above plot the 30% (dark red), 60% (red) and 90% (light red) confidence intervals for *tertiary education attainment* forecast constructed using the best econometric specification presented in section 4.1. More certainty would mean wider forecast range. Based on these confidence intervals we have constructed the qualitative forecasts presented in section 5. The blue points represent the forecasts obtained from the birth-cohort model described in section 6.

Please note that the scale of the vertical axis might be different depending on the country.







Note: The shaded areas in the figures above plot the 30% (dark red), 60% (red) and 90% (light red) confidence intervals for *tertiary education attainment* forecast constructed using the best econometric specification presented in section 4.1. More certainty would mean wider forecast range. Based on these confidence intervals we have constructed the qualitative forecasts presented in section 5. The blue points represent the forecasts obtained from the birth-cohort model described in section 6.
Please note that the scale of the vertical axis might be different depending on the country.

Annex 3. Alternative econometric specifications for the twofold Europe 2020 headline target

Despite having good theoretical support, we did not include mortality risk in any specification. Different proxies were tried but none was found significant; as a first look it appears that these indicators are changing too slowly over time. From an empirical perspective, this is in line with the evidence provided in Acemoglu and Johnson (2007) and Hazan (2009) who find no evidence that increasing life expectancy has caused a rise in schooling over the last century.

All the variables in the table below were used as proxies for mortality risk which appears as a determinant of education attainment in theoretical models. The comparative statistics (all EU27 member states) from Table A1 provide an illustration of the standard variation differential between ET2020 headline targets and mortality risk proxies.

Table A.1

Variable name (change in log of …)	**Number of observations**	**Mean**	**Standard deviation**	**Min**	**Max**
Early school leavers	*241*	*-3.3*	*10.0*	*-55.4*	*38.5*
Tertiary education completion	*241*	*5.2*	*7.3*	*-31.1*	*59.6*
Life expectancy at birth	*241*	*0.4*	*0.4*	*-1.0*	*1.8*
Mother's age at childbirth	*237*	*0.5*	*0.5*	*-3.5*	*3.9*
Young dependency ratio 1st variant	*240*	*-1.2*	*1.6*	*-5.3*	*3.1*
Age dependency ratio 1st variant	*240*	*-0.1*	*1.1*	*-3.7*	*3.4*
Fertility rate	*237*	*0.9*	*3.3*	*-11.3*	*13.6*

Source: Eurostat and authors' calculation for EU27 member states over 2003-2011 period

Some other theoretical and empirical approaches pertaining to labour economics, deal with education and its relation with physical investment or population migration (see Acemoglu 1996, Eggert et al. 2010). We tried to include variables such as investment, capital share and migration into our models but they were not robust and were discarded.

Nevertheless, as already explained in section 4, some of these omitted factors could still have an influence on the education attainment developments in some particular countries. Our econometric models are just simple representation of the reality. They were not designed to offer a complete and comprehensive picture about the mechanisms involve in schooling decisions. Regarding the expected dynamics of some of these potential important factors, country specific judgements can be added to improve the message of the forecast and therefore alter our qualitative evaluation.

Following, we present alternative econometric specifications for both benchmarks to highlight the robustness of our empirical findings.

Table A.2. Different estimated specifications *early school leavers* benchmark

Δ log(early school leavers)	OLS estim.	LSDV estim.	OLS estim.	OLS estim.	OLS estim.
Δ log(females' share low education, 35-44) t	0.13***	0.17***			
	(0.05)	(0.05)			
Δ log(males' share low education, 45-54) t	0.23***	0.20**			
	(0.07)	(0.08)			
Δ log(adults' share low education, 35-44) t			0.12**	0.24***	
			(0.06)	(0.07)	
Δ log(adults' share low education, 45-54) t			0.26***	0.34***	0.46***
			(0.09)	(0.09)	(0.10)
Δ log(unemployment rate) t-6	-0.07**			-0.07**	
	(0.03)			(0.03)	
Δ log(unemployment rate) t-7		-0.08***			
		(0.03)			
log(output gap) t-4			0.44***		
			(0.12)		
Δ log(unemployment rate differential by skill ISCED 3-4 vs. 0-2) t-5					-0.03*
					(0.01)
Constant	-1.82***	-1.35	-1.64***	-0.61	-0.10
	(0.55)	(1.64)	(0.62)	(0.69)	(0.81)
Observations	276	276	295	278	184
R^2	0.137	0.198	0.150	0.174	0.175
AIC	1815.78	1831.68	1965.07	1853.17	1283.11
BIC	1859.22	1940.30	2001.95	1903.96	1295.97
No. countries ^)	19	19	20	19	21
Time dummies +)	Yes	Yes	Yes	Yes	No
Country dummies	No	Yes	No	No	No
Out-of-sample forecasting properties:					
RMSE 1 year ahead ++)	1.10	1.12	1.13	1.13	1.10
RMSE 2 years ahead	1.77	1.90	1.84	1.84	1.88
RMSE 3 years ahead	2.47	2.81	2.66	2.67	2.52
RMSE 4 years ahead	2.77	3.26	3.07	3.09	2.85
Wooldridge test for autocorrelation ++)	Ok	Ok	Ok	No	Ok
Arellano-Bond test for higher order autocorrelation ++)	Ok	No	Ok	Ok	Ok
Baltagi-Wu LBI test	-	Ok	-	-	-
Coefficients stability over different estimation samples	Ok	Ok	No	Ok	No

Standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

^) We start with the full set of EU27 countries and then excluded some of them according to the weights assigned in a robust regression specification. We use mainly unbalanced panels for early school leavers, although some specifications did include balanced panels.

+) We use time dummies only for the period before 2000 in order to control for the fact that some panels were unbalanced.

++) These tests were computed on different samples of different sizes. The samples start with the first year of the estimation and end in each of the following years 2006-2012. Accordingly, our RMSE indicators were computed on the same estimated specification, but over expanding time samples. For the autocorrelation tests, "no" means that the tests were not robust in rejecting autocorrelation, while "ok" means vice-versa.

Note: As an additional check we estimated the same specifications correcting for unobserved heteroskedasticity using the robust standard error option in Stata, but the statistical significance did not change much.

Although the business cycles are hardly influenced by the schooling decisions of very young individuals, please note that there could be a slight endogeneity issue associated with the specification including the output gap. Various lags were tried without altering the conclusions; the specification presented here was nevertheless the best one according to our selection criteria.

The relation between early school leavers and youth unemployment was also investigated, but we were not able to find specifications with longer enough lags to mitigate the obvious endogeneity problem evident in this case, i.e. that youth unemployment is a '*result*' (from an econometric perspective) of education decisions.

Table A.3. Different estimated specifications *tertiary education attainment* benchmark

Δ log(tertiary education attainment)	OLS estim.	LSDV estim.	OLS estim.
Δ log(females' share high education, 55-64) t		0.19***	
		(0.05)	
Δ log(adults' share high education, 55-64) t	0.34***		0.36***
	(0.06)		(0.06)
Δ log(labour productivity) t-13	0.58**		
	(0.28)		
Δ log(TFP) t-12			0.40*
			(0.21)
Δ log(employment) t-11		-0.36*	
		(0.19)	
Constant	0.48	4.56***	1.37
	(0.60)	(1.16)	(1.10)
Observations	144	132	132
R^2	0.322	0.345	0.392
AIC	776.26	712.82	693.35
BIC	791.11	753.18	733.71
No. countries^)	12	11	11
Time dummies	Yes	No	Yes
Country dummies	No	Yes	No
Out-of sample forecasting properties:			
RMSE 1 year ahead +)	1.38	1.21	1.32
RMSE 2 years ahead	2.23	1.92	2.31
RMSE 3 years ahead	2.94	2.46	3.10
RMSE 4 years ahead	3.50	2.65	3.20
Wooldridge test for autocorrelation +)	Ok	Ok	Ok
Arellano-Bond test for higher order autocorrelation +)	Ok	No	Ok
Baltagi-Wu LBI test	-	Ok	-
Coefficients stability over different estimation samples	Ok	No	No

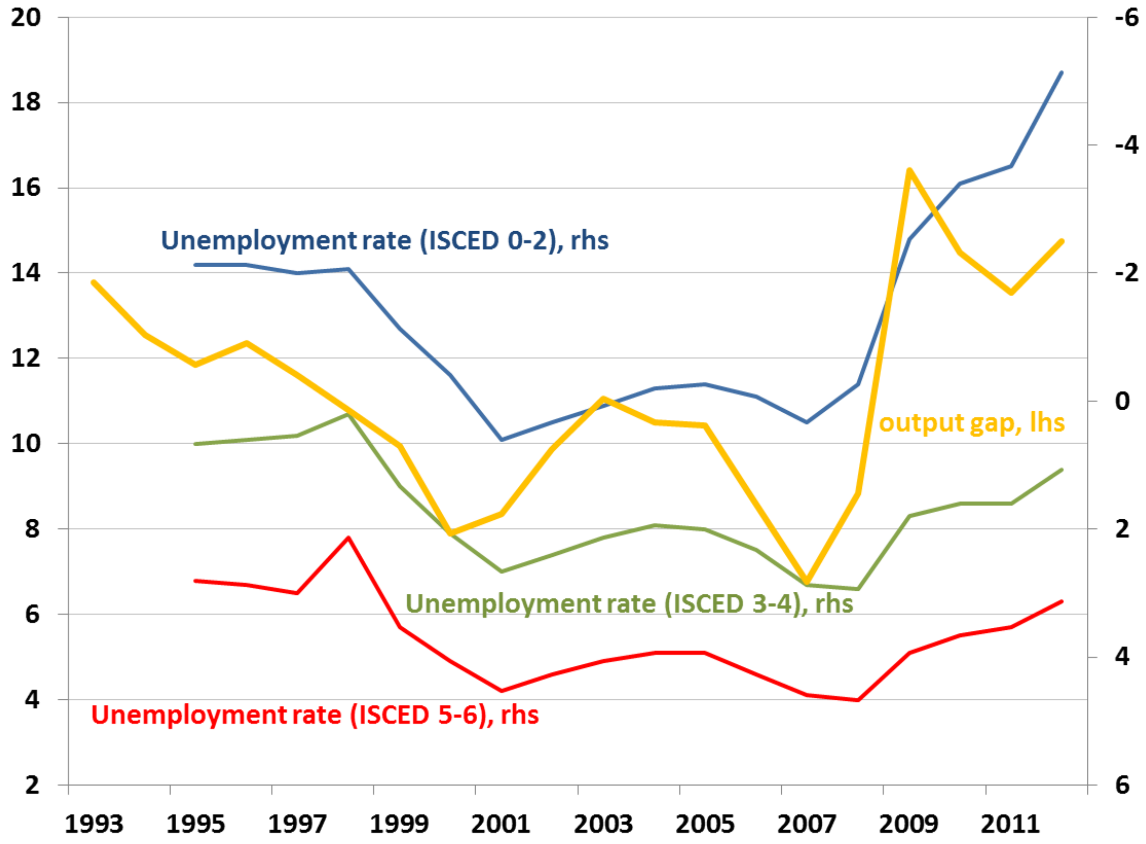
Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

^) First, we select member states with long enough data for the specifications illustrated here. Then we start excluding countries according to the weights assigned in a robust regression specification. We prefer working with balanced panels as their statistical properties were better than for unbalanced specifications.

+) These tests were computed on different samples of different sizes. The samples start with the first year of the estimation and end in each of the following years 2006-2012. Accordingly, our RMSE indicators were computed on the same estimated specification, but over expanding time samples. For the autocorrelation tests, “no” means that the tests were not robust in rejecting autocorrelation, while “ok” means vice-versa.

Note: As an additional check we estimated the same specifications correcting for unobserved heteroskedasticity using the robust standard error option in Stata, but the statistical significance did not change much.

Annex 4. Unemployment rate (%) by skills and the economic cycle (output gap in % of potential GDP)



Source: Eurostat and AMECO, data for EU15

Note: We use EU15 data instead of EU27 because we wanted to have a longer time sample for a better illustration of the correlation.

Annex 5.
Assumptions behind the projections for the education attainment of adult population, by age groups

Parents' education attainment is extrapolated up to 2020 using a simple rule, illustrated in the equation (1) below. It may not be the best one for each country, but it is a good compromise between simplicity and efficiency. Ignoring inward and outward migration or mortality risk differential between individuals with different education and assuming that *all the education decisions were taken in the past* given the age groups we are considering, we have:

forecast for: ***Pop*** *(education=**E**) [age group=**A**]* ten years ahead =

Pop *(education=**E**) [age group=**A-10**]* today + λ * *cohorts'gap* (1)

where Pop is the share of individuals from the total population with a maximum education attainment level E , while the age groups *[A]* and *[A-10]* would refer to different population cohorts, which can vary depending on the chosen econometric specification (see Table 2 in the main text for more details about the age groups). Ignoring the last term in the equation above, we can say that after a given time period we expect the same, but older, individuals to be counted in the indicator.

The gap between consecutive cohorts could be significant in some countries and suggests that our assumptions about migration, mortality and education decisions might be an over simplification of the true developments. The coefficient λ is a smoothing adjustment coefficient with a value set at 0.5 that was added to narrow the gap in education levels between consecutive cohorts[46], notation *cohorts'gap*. This calibration means that the gap would be closing annually half of its previous year value.

Table bellows summarize the projections for parents' education used in this forecasting exercise based on the equation (1) above:

Table A.4

	share of females' with low education, age 35-44		***share of males' with low education, age 45-54***		***share of adults with high education, age 55-64***	
	2012	**2020[f]**	**2012**	**2020[f]**	**2012**	**2020[f]**
AT	17.1	13.1	11.0	10.2	16.7	18.6
BE	19.8	16.3	32.4	23.6	25.3	30.9
BG	15.6	4.4	17.2	16.4	21.5	22.3
CY	16.6	14.7	21.5	16.2	24.1	29.9
CZ	5.4	10.6	4.6	3.6	12.6	15.9
DE	14.7	13.8	10.7	11.7	26.4	26.6
DK	15.8	17.1	22.9	20.2	28.7	31.2
EE	7.3	9.3	6.9	12.0	35.6	38.5
ES	34.8	30.1	49.4	43.6	19.0	25.6
FI	7.1	7.2	15.7	14.0	31.4	38.9
FR	19.6	15.2	29.9	23.1	19.6	21.8

[46] Some countries, such as UK and Luxemburg for example, have big gaps in the education attainment for two consecutive cohorts, probably due to high migration. Migration flows were not included in this study for lack of robustness in the panel econometric specifications. Alternative scenarios could nevertheless consider different projections for parent' education rather than those included here, depending on the assumptions behind migration flows.
For countries where the methodological breaks in education attainment data appearing after 2003 are explaining a large part of the *cohorts'gap*, we use a slightly higher smoothing adjustment coefficients in the rage [0.8, 0.9]. This is justified by the fact that a break in 2007 for example, would mean that data before 2006 for younger cohorts (which will translate into forecasts for the period 2013-2016 for older cohorts) would belong to the old methodology and could be therefore slightly inadequate. We do not rule out other alternative scenarios or extensions of the analysis in future versions of the paper.

EL	23.9	18.7	39.1	33.2	19.5	22.2
HU	16.8	13.4	14.9	14.8	15.4	18.3
IE	17.0	12.4	33.6	24.2	24.9	30.2
IT	34.8	25.8	49.3	43.7	11.4	12.0
LT	5.2	9.1	3.2	6.0	25.5	27.5
LU	21.9	15.3	21.6	16.9	26.4	27.9
LV	6.7	12.3	7.9	11.4	22.0	24.3
MT	57.9	48.0	62.3	53.9	7.8	9.3
NL	20.5	15.3	26.9	23.7	26.8	29.8
PL	6.7	5.0	9.5	8.9	12.6	15.1
PT	49.9	40.8	75.4	71.0	11.1	10.4
RO	19.0	25.0	16.7	16.4	9.5	10.8
SE	12.0	10.9	17.8	13.6	28.7	29.3
SI	10.7	4.6	15.5	14.7	17.2	20.0
SK	6.2	5.7	6.6	5.9	13.7	14.9
UK	19.5	16.0	24.0	21.6	30.8	33.0

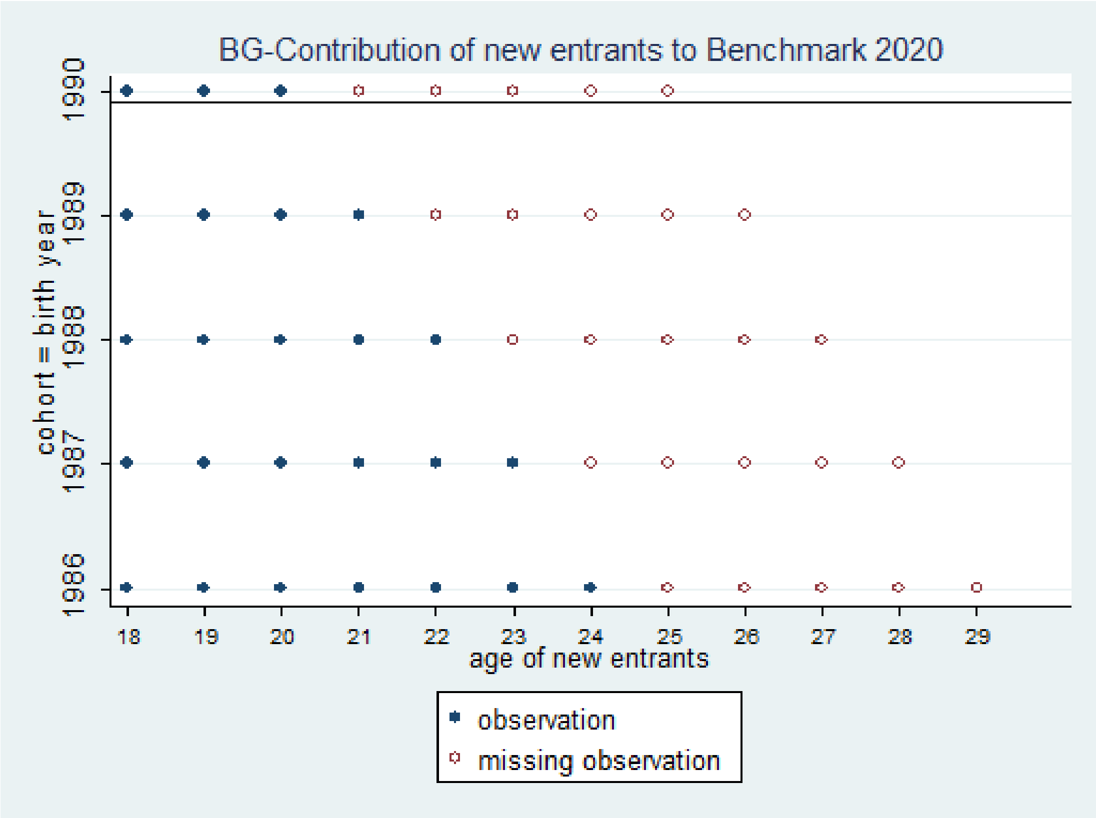
Note: Data for 2012 are from Eurostat while data for 2020 are authors' calculations

Annex 6. Additional details about the birth cohort model

Step 4: Computing the total population aged 30 to 34 for each year of the target

For each year of the target, we calculated the number of individuals belonging to the cohorts that are counted into the benchmark. For example, for the benchmark value 2020, we compute the total population of individuals born between 1977 and 1990. For example, the cohorts included in the benchmark calculation for the year 2020, are depicted in chart below.

Contribution to Target 2020 – Cohorts



Data on the total population aged 30 to 34 for each year of the target is taken from Eurostat. In particular, Eurostat provides information on the population by individual ages born in a certain year (i.e. when they are 18, 19, 20, ..., 29 years old). This implies that we have multiple observations for a specific birth cohort. One example for Bulgaria is presented in Table A.5 below for the cohort year 1974. Information on the population of this age cohort in the year 1998 (people are 24 years old) until the year 2003 (people are then 29 years old) is displayed. Moving from 1998 to 2003 the population of this birth cohort reduces from 118 295 to 113 375. The population of the birth cohort is finally calculated by taking the average sample value of the birth cohort population[47].

[47] An alternative would be to only consider the most recent data on total population since these are then the people to be counted into the calculation of the headline target.

Table A.5: Example of population of birth cohorts for BG

Population	Cohort	Year	Age
118 295	1974	1998	24
118 178	1974	1999	25
118 079	1974	2000	26
117 924	1974	2001	27
113 479	1974	2002	28
113 375	1974	2003	29

European Commission
EUR 26077 EN – Joint Research Centre – Institute for the Protection and Security of the Citizen

Title: Forecasting the Europe 2020 headline target on education and training.

Authors: Catalin Dragomirescu-Gaina, Anke Weber

Luxembourg: Publications Office of the European Union

2013 – 42 pp. – 21.0 x 29.7 cm

EUR – Scientific and Technical Research series – ISSN 1831-9424 (online), ISSN 1018-5593 (print)

ISBN 978-92-79-32509-0 (pdf)
ISBN 978-92-79-32510-6 (print)

doi:10.2788/92332

Abstract

This analysis aims at proposing simple econometric models that can be used to forecast early leavers from education and training and tertiary education attainment benchmarks up to year 2020. The models are built on the theoretical framework of human capital and optimal schooling decisions and then estimated in a panel setting to better deal with a limited dataset. By looking back at the time period of enrolment and graduation, our approach could be seen as an attempt to identify the determinants that shape the education decisions of young individuals.

We construct the forecasts under very simple assumptions about the expected adults' education attainment and given the determinants of schooling decisions uncovered by our empirical analysis. The forecasts tell us how early school leaving and tertiary education attainment are likely to develop over the next years if nothing changes in terms of policy measures. This very strong assumption provides scope for policy action especially for those countries where the expected developments of model's determinants are not enough to foresee a positive outcome.

LB-NA-26077-EN-N

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 standards, methods and tools, and sharing and transferring its know-how to the Member States and international community.

Key policy areas include: environment and climate change; energy and transport; agriculture and food security; health and consumer protection; information society and digital agenda; safety and security including nuclear; all supported through a cross-cutting and multi-disciplinary approach.

Publications Office

ISBN 978-92-79-32509-0

9 789279 325090