

Social Investments for resilient economies

Final working document

**Informal Working Group
Social Investment**

Contents

Introduction	1
Social investments and their expected support to upward economic and social convergence and resilience	3
Social Investment: what in a name	3
The idea of Social Investment	3
Social investment in a life-course perspective	4
Evidence of social investments and social protection returns	7
Human capital as engine of economic growth	9
Social protection: economic stabilization and allocative efficiency	10
Social investment employment, poverty, inequality and health returns	14
Employment effects of complementary social investment policies	14
Poverty reduction effects of complementary social investment policies	16
Health outcomes	18
Key methodologies to monitor and evaluate social investments	21
Monitoring arrangements, indicators and the importance of data infrastructure	21
Data to evaluate social investments: state-of-play and way forward	23
Overall approach to evaluate the return on social investments	25
Overview of the approaches and methods to measuring social investments returns	27
Ex ante micro approaches	30
Ex-post micro evaluation	31
Randomized Control Trials	32
Matching methods	34
Regression Discontinuity Design (RDD)	35
Differences in Differences (DiD)	35
Qualitative methods for evaluation	36
Ex post macro statistical approaches	37
Distributional impact assessments (DIA)	40
Distributional impact assessment and budgetary plans	47
Distributional impact assessment, data used and implementation	49

Introduction

On 26 April 2023 the European Commission presented a proposal for a new economic governance framework. The objective is to strengthen debt sustainability and promote sustainable and inclusive growth. According to the proposal, I am v.

The European Pillar of Social Rights figures prominently among the core EU common priorities which the new fiscal framework is expected to promote and is well integrated in the European Semester. This prominence does not come as a surprise. In a context of multiple crises, well-functioning and inclusive social investment welfare states are a key component of a socially and economically resilient society. Looking back on the long decade since the Great Recession, it is undeniable that, far from 'crowding out' scarce resources, well-funded and active welfare states have proved to be a condition sine qua non for resilient economies.

As highlighted in the Commission Strategic Foresight reports¹, enhancing the socio-economic resilience of a society entails multiple actions, from a sound fiscal position to be able to absorb shocks, adaptation as the time of exposure to shock and its intensity increases, to transformation to ensure that the system not only avoids collapses but also finds its new sustainable development path (bouncing forward)².

In times of crises, inclusive and adequate social protection schemes turn out to be the most important feature in predicting the absorptive capacity of a country. It may be more efficient to address an unforeseen economic shock like a pandemic with an inclusive welfare state that provides strong automatic stabilizers, such as broad and well-organized access to sickness and unemployment benefits and short-time working arrangements. When focusing on the medium to long-term and on the capacity of countries to 'bounce forward', however, what becomes critical are well-designed and targeted active labour market policies, as well as up- and reskilling measures and inclusive early childhood and long-term care policies, which have a positive impact on potential economic growth and employment.

In recent years, the notion of social investments has gained purchase as a policy compass for recalibrating the welfare state and the need to be competitive in the new knowledge-based society³. The objective is not only to protect from social risks ex post, but to enhance individuals' opportunities and capabilities to address ex ante social risks typical of post-industrial economies, while ensuring the high (quality) employment needed to sustain the fiscal 'carrying capacity' of the welfare state, which relies on the number in employment and their productivity.

At the core of the social investment paradigm sits the idea of **life-course multiplier**. Over the life-course, social investments reap wellbeing returns in a virtuous circle, in terms of employment opportunities, well-being, and gender equity and the mitigation of intra- and intergenerational poverty. The life-course multiplier features prominently in the recent report by the High-level Group on the future of social protection and the welfare state in the EU⁴. At the micro-level of individuals and households, it suggests how social investments, from early childhood, improve material wellbeing (employment and income) and help mitigate social risks later in life, through opportunities for skills acquisition, the easing of (gendered) labour-market transitions, and healthy

¹ Available here: https://commission.europa.eu/strategy-and-policy/strategic-planning/strategic-foresight_en.

² Giovannini, Benczur, Campolongo, Cariboni, Manca (2020) Time for transformative resilience: the COVID-19 emergency.

³ Vandenbroucke, F., Hemerijck, A. and Palier, B. (2011), "The EU Needs a Social Investment Pact", OSE Paper Series, Opinion paper n°. 5 (pag 22-23).

⁴ Available here: <https://op.europa.eu/en/publication-detail/-/publication/1b51a40b-591e-11ee-9220-01aa75ed71a1/language-en/format-PDF/source-294385988>

ageing. At the macro-level, the multiplier entails a 'double dividend', of greater and more gender-balanced employment and productivity gains supporting fair and adequate as well as sustainable social protection.

While the reference to the Pillar in the April Commission proposal for a new economic governance framework is positive, the criteria for deciding which investments and reforms are eligible for an extension of the debt-adjustment plans has not yet been defined. Such criteria will depend on the outcome of the ongoing negotiations on the EGR. Based on the Commission proposal of April 2023, an extension of the adjustment period can be requested by Member States if they implement reforms and investments that would support **economic growth and fiscal sustainability** while being aligned with the EU priorities and respond to the main challenges identified in the Country Specific Recommendations. At the same time, the proposed 2023/0138 (COD) Regulation specifies that the plans **shall explain how they will ensure the delivery of a relevant set of reforms and investments**. Each of the reform and investment **shall be sufficiently detailed, front-loaded, time-bound and verifiable**.

Against this background, the Spanish Presidency and the Belgian Presidency of the EU have launched an Informal Working Group on Social Investment (IWGSI). The main objective of the IWG is – without interfering with the ongoing negotiations on the Economic Governance Review (EGR) – to provide input based on empirical evidence regarding micro and macroeconomic returns of social investments and reforms as well as monitoring and evaluation methodologies for tracking social investment returns, to support a political debate.

The purpose of this note is to back the reflection of the IWGSI in these two directions.

The note has been completed on the 11th March 2024, following the 9 meetings of the Informal Working Group on Social Investment. The text has been written by the chairs of the IWGSI Dr. Francesco Corti and Dr. Milagros Paniagua, with the key support of the teams of the Belgian and Spanish Presidencies of the Council of the EU. A special thank goes in particular to Rudi Van Dam, Isalyne Stassart, Deniz Demiral, Pedro Eduardo Bernard Silva, Javier Lopez Somoza, Agustin Berasaluce Astarloa, Alberto Barranco Garcia, Lucia Lorenzo Meyer and Monica Martinez Bravo.

Brussels, 30 June 2024

Francesco Corti and Milagros Paniagua

Social investments and their expected support to upward economic and social convergence and resilience

Social Investment: what in a name

Over the past two decades, academic experts from various disciplines have started to rethink the interaction between economic progress and social policy: from equity-efficiency ‘trade-offs’ (Okun, 1975) and associated ‘trilemmas’ (Iversen and Wren, 1998) to Pareto-optimal “social investment synergies”, even in Rawlsian terms of benefiting the worst off (Esping-Andersen et al, 2002; Esping-Andersen, 2009; Morel et al, 2012; Hemerijck, 2013; 2017; Hemerijck et al., 2016). Social investment is today understood as welfare provision that helps ‘prepare’ individuals, families and societies ex ante to respond to the changing nature of social risks in advanced economies, by investing in, upkeeping and protecting human capabilities from early childhood to old-age, at the same time as pursuing policies that ‘repair’ social misfortune ex post, after moments of economic and personal crisis.

The idea of Social Investment

The term social investment was conceptualized for the first time by Anthony Giddens in 1998, as an attempt to redefine social policies’ role in the new post-industrial globalized and knowledge-based society. An explicit impetus to the policy theory behind the social investment approach was given with the publication of the book, edited by Gøsta Esping-Andersen, *Why We Need a New Welfare State* (Esping-Andersen et al. 2002). Today, social investment is widely embraced in policy reports and documents of the European Union (European Commission, 2013), the OECD (2008; 2014), and the World Bank (2016) as a compass for ‘inclusive and sustainable growth’, able to inspire actors from across the political spectrum. The High-Level Group report on *The Future of Social Protection and the Welfare State in the EU* (2023) for the European Commission is surely the most assertive official endorsement of social investment to date.

At the core of the new social investment paradigm stands the idea of **economic return** and **measurable welfare provisions**: social policy as a productive factor. The problem that many States were experiencing starting from the mid-Seventies and especially in the Nineties, was that of persisting high public debt, high level of unemployment, low growth and continuous demographic ageing, which indeed posed significant difficulties both on the carrying-capacity of the welfare state and individuals’ life-course transitions. Social spending became a big financial burden for countries which had to cope with exogenous economic constraints while paying welfare benefits to rapidly ageing populations (Ronchi 2018). For this reasons, the initial aim of social investment was to make people’s working lives longer and to extend the tax-payers base, by guaranteeing workers proper education, longer lifelong learning and long-term care.

The overarching objective of social investment welfare provision is to enhance people’s opportunities and capabilities to resolve social risks typical of post-industrial societies ex-ante. Early childhood education and care, vocational training over the life-course, capacitating active labour market policies, work-life balance policies like paid parental leaves, lifelong learning, and long-term care; what all these policies have in common is that they transcend – but not replace – the compensatory rationale of post-war social security that protected (predominantly male) workers and their (stable) families against industrial risks ex-post (Hemerijck, 2018). In other words, the notion of social investment shifts the terms of the welfare state debate from the ageing quandary of rising pressures on public spending, closely associated with population ageing,

toward a more comprehensive understanding of how welfare provision interacts with demography, including family formation, education, skills, labour supply and productivity.

Social investment thus understood boils down to the hypothetical conjecture that a well-organized welfare state is an asset, and not a liability to the economy, as it contributes to productive capacity of the economy and its stability in the face of adverse economic conditions. The **economic logic of social investment** explicitly focuses on increasing the number (quantity) and productivity (quality) of current and future employees – or ‘the carrying capacity’ of popular income protection and social security. Following the economist Richard Musgrave (1989) and the sociologist John Myles (2002), we can conceptualize the carrying capacity according to a cost-benefit equation, where social investment shifts attention from the *numerator* side of welfare state costs to the *denominator* side concerning the quality and quantity of those carrying the costs:

$$\frac{\text{Number of welfare recipients}}{\text{Number of paid workers}} \times \frac{\text{Average consumption of welfare recipients}}{\text{Average productivity of paid workers}} = \frac{\text{Welfare spending}}{\text{Tax base}}$$

The work-income-family nexus is *lynchpin* of social investment welfare provision. The greatest contribution to the carrying capacity, and indeed the key driver of the social investment orientation in the welfare state in recent decades, has been the rise of women entering the labour market ever since the 1990s. Like demographic ageing, this evolution represents a major feat of societal emancipation and progress, but at the same time a success formula that must be managed in terms of accommodating welfare provision to balance out work and family care. Female scholars, including Barbara Hobson (1990), Jane Lewis (1992), Ann Orloff (1993), and Diane Sainsbury (1996), were among the first to highlight the new set of gendered challenges to the welfare state, by stressing the need for childcare, long-term care, parental leave, and targeted support for lone parents, mostly women, as part and parcel of post-industrial welfare provision.

Social investment in a life-course perspective

At the core of 21st century welfare states is the **life course** perspective. Indeed, throughout the life course all citizens rely on all kind of welfare provision for childcare, education, social insurance, retraining and activation, health and long-term care, pensions and family support. As observed above, the tertiarisation of the labour market, the increasing depletion of skills due to rapid technological change and the progressive erosion of the ‘standard employment relationship’ call on welfare states to rethink labour market policy and employment regulation with the aim of minimizing precarious employment and underemployment. At the same time, the increasing participation of women to the labour market and the changing family structure require a reinterpretation of the welfare provisions, with special attention placed on avoiding career interruptions for women with small children, promoting dual earner families, with a gender-equal parental leave. Overall, since lives are longer, more diverse and volatile, it is important for the welfare state to guarantee smooth transitions in life course, for instance, when people move from education into their first job, when they aspire to have children, when they experience spells of labour market inactivity, and, finally, when they move to retirement.

Modern welfare states are therefore a balanced mix of two dimensions: social investment and social protection.

- **Social investment** includes policies aimed at raising and upkeeping the quality of the

'stock' of human capital and capabilities, as well as easing and improving the 'flow' of contemporary labour-market and (gendered) life-course transitions.

- **Social protection** refers to policies targeted at maintaining strong universal safety net 'buffers' for micro-level income protection and macro-economic stabilization (Hemerijck 2014; 2015; 2017).

More in details, being directed towards capacitating interventions, the stock function of modern welfare states aims at enchainning and maintaining human capital or capabilities over the life-course in ageing society, by bringing under one roof adjustable bundles of professional assistance from child- to elderly care, including skill enhancement and training services in case of unemployment, health, family and housing support (Sabel et al., 2017). In so doing, high-quality systems of vocational training, education, retraining programs and lifelong learning arrangements best ensure long-term employability and high employment participation (Hemerijck 2013). Moreover, given that the children make up the future workforce, investing in better education and affordable childcare will ultimately lead to higher levels of productivity and employment, when ageing predicament reaches its pinnacle.

Arguably, the 'flow' function is about efficient and optimal allocation of labour and employment over the lifespan, making sure that school-leavers, parents (especially mothers), unemployed workers, older workers or the persons with disabilities can return to work as fast as possible through active labour market policies, job matching, work-life balance services and assistance during vulnerable transitions. The flow function is to be understood in terms of helping to bridge critical life-course transitions. This, however, should not pave the way with maximum labour market flexibility. On the contrary, as stressed by Guenther Schmid (2015), the social investment turn is characterized by a shift from the mantra of "making work pay" to "making transitions pay", through the provision of 'active securities' or 'social bridges' across volatile transitions between jobs, thereby ensuring sustainable and longer working careers and, by implication, adequate pensions after retirement.

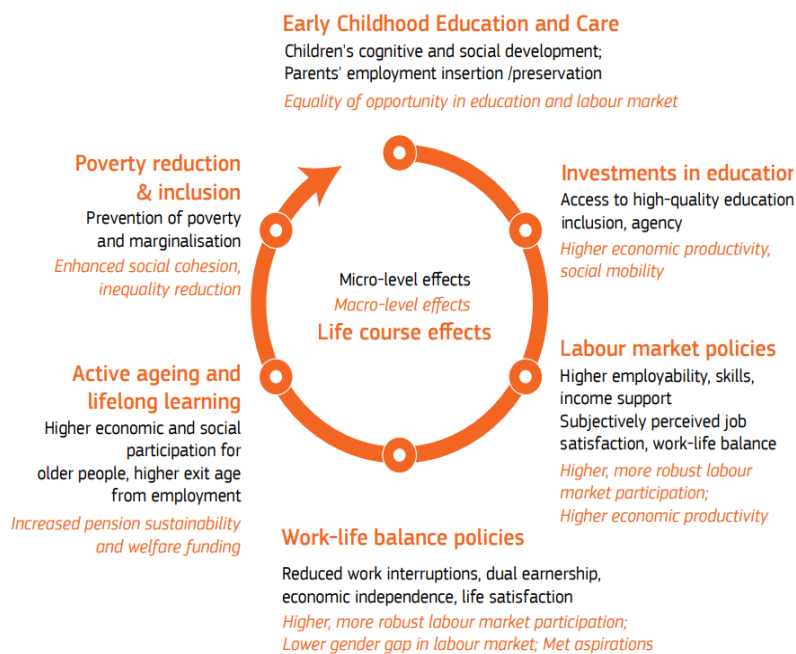
Finally, the 'buffer' function is about securing adequate and inclusive income protection, thereby further stabilising the business cycle and buffering economic shocks. A traditional example of this in case is unemployment insurance schemes. Adequate income protection is a precondition for an effective social investment strategy, as income buffers help to compensate and mitigate social insecurity at the micro level and thus provide the necessary financial security for people to develop their human capital while at the same time securing economic stabilization at the macro level.

The three welfare functions are not to be considered as silos. By contrast, in everyday policy-practice, there is ample overlap between the policy functions of 'stocks', 'flows' and 'buffers'. They can run as 'institutional complementarities' with the operation of one social investment function being enhanced by one of the others or both in a mutually reinforcing fashion (Hall and Soskice, 2001). Policy provisions that seemingly focus on one of the three functions often back up the others in an interconnected fashion and need to do so (Dräbing and Nelson 2017). For example, childcare services act both to provide children with education (stock function) and to guarantee proper work-life balance for families and a speed return to work especially for women (flow function). Moreover, if designed in an inclusive way, childcare services can play a fundamental role in eradicating child poverty and breaking the cycle of inter-generational transmission of disadvantage (buffer function). By contrast, as stressed by Van Lancker (2013), isolated social investment policy innovations, without taking into account institutional complementarities, can be counterproductive and extremely costly. The example in case is, again, childcare policies.

The growing evidence on how effective combined welfare policies reinforcing the proficiency of each other has allowed Anton Hemerijck (2015; 2016; 2017) to conjecture the operation of a social investment '**life-course multiplier**' effect, whereby high quality early childhood care over time

contributes to higher levels of educational attainment, which in turn, together with more tailor-made vocational training, can spill over into higher and more productive employment in the medium term. To the extent that employment participation is furthermore supported by effective work-life balance policies, including adequately funded and publicly available childcare, higher levels of (female) employment with lower gender pay and employment gaps and ultimately pension gaps can be foreseen. More opportunities for women – and men – to combine parenting with paid labour or training is, in addition, likely to have a dampening effect on the so-called ‘child gap’, the difference between the desired number of children per couple (aspirational fertility) and the actual number of children (realized fertility) (Bernardi 2005). A final knock-on effect is a higher effective retirement age, provided through the availability of active ageing policies, including portable and flexible pensions, for older cohorts. Such multiplier effect of institutional complementary social policies has then an intergenerational effect, whereby return are transmitted across generations. If children are cognitively stimulated and do not start out their lives in poverty, they are less likely to be poor in adulthood and more likely to be employed longer and in productive jobs, with the effect that their offspring will not grow up in poverty, which then gives an extra boost to tax revenue and thereby a stronger funding base for costly but beneficial investment (Hemerijck et al. 2016).

Figure 1. The Social Investment life-course multiplier

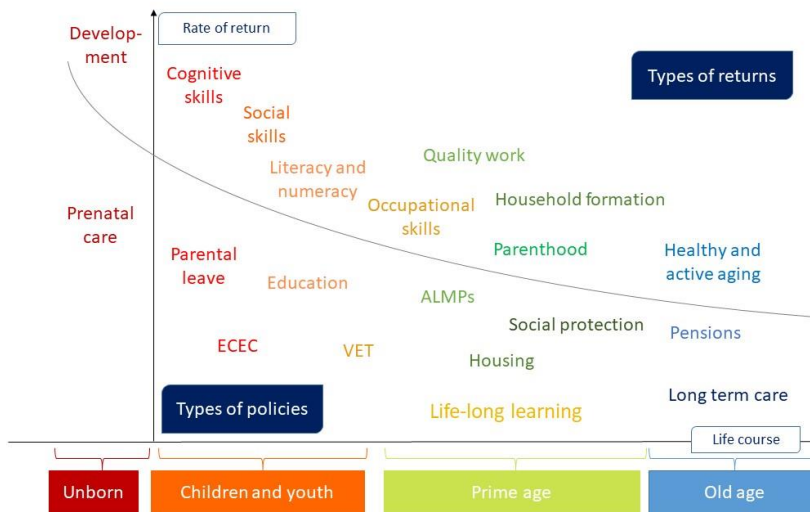


Source: Source: A. Hemerijck, S. Ronchi, I.Plavgo, Social investment as a conceptual framework for analysing well-being returns and reforms in 21st century welfare states, Socio-Economic Review, 2022

As put in the ESDE (2019) report⁵, at young ages, the returns tend to be highest for children from disadvantaged backgrounds, implying that such investments can promote both efficiency and equity from a longer-term perspective.

⁵ Add reference to ESDE 2019... [Publications catalogue - Employment, Social Affairs & Inclusion - European Commission \(europa.eu\)](https://publications.ec.europa.eu/publications-catalogue)

Figure 2. Expected returns on social investment and rate of return, by life stage. Returns on social investment are particularly high at early life stages



Source – ESDE 2019. [Publications catalogue - Employment, Social Affairs & Inclusion - European Commission \(europa.eu\)](https://ec.europa.eu/europeaid/publications-catalogue)...

Evidence of social investments and social protection returns

As stressed above the idea of social investment has been accompanied by the debate on **measurable returns**. Such returns are materializing in terms of micro- and macro- economic and social outcomes. Among all possible social investment policies, education, activation and health policies are generally found to provide high economic and social returns in the medium to long-term through enhancement of the stock and quality of **human capital**. At the **micro-level**, education is traditionally a predictor of higher individual resilience. High-quality early childhood education and care in the long term is strongly correlated with better educational attainment, higher labour market participation, and reduced economic inequality, especially for children from disadvantaged socioeconomic background. At the **macro level**, the role of human capital formation in fostering economic growth in the medium to long term is by now widely recognized. Both reforms and investments in quality education (from an early age) and training (up- and re-skilling) are of the utmost importance in this sense. Such measures are found to directly foster economic growth and competitiveness, increase productivity, including via stronger innovative capacity and absorption of new technologies, and also by ensuring a better matching between skills demand and supply. Well-functioning and effective active labour market policies (ALMP) and public employment services provide also both individual economic and social returns in the medium to long term by increasing the efficiency of the labour market, mitigating skills mismatches and shortages, and strengthening labour market participation of under-represented groups. At the micro-level, effective ALMP programs reduce the cost and time required for a worker to be matched to a job and improve the quality and duration of such a match.

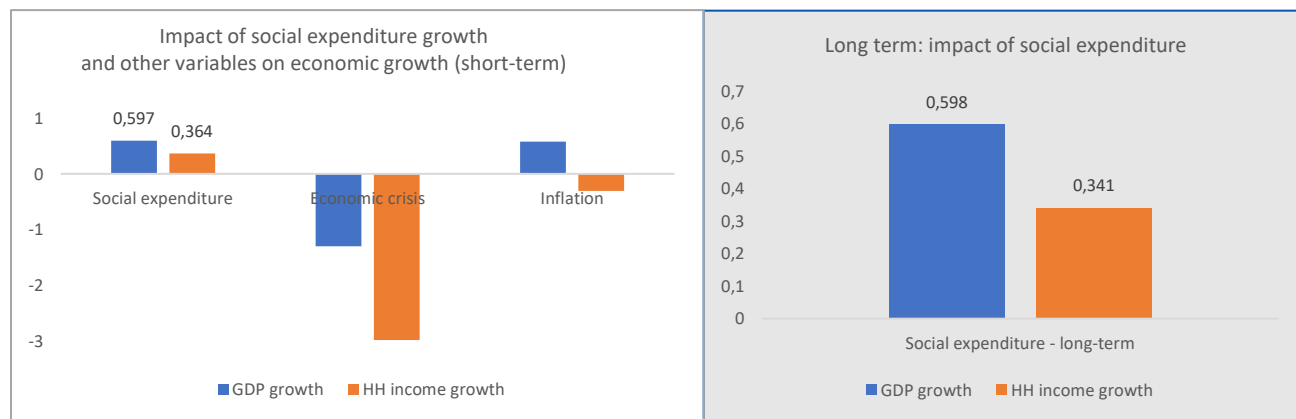
While educational and activation policies materialize their micro- and macro returns in the medium to long term **social protection policies** (e.g. typically unemployment benefits or other income support) are conducive to macro-economic stabilization in the short term. At the **macro level**, impact resilience, i.e., the country capacity to absorb a shock, is closely associated with high government expenditures on social protection when the latter is accompanied by inclusive

schemes which also secure the most vulnerable groups. Social protection has indeed a significant and positive impact on growth via the income and consumption channel. At the **micro individual level** social protection is complementary to social investment, since for example, without adequate social benefits, the provision of childcare services would not address the needs of the most vulnerable children and their families.

A recent ECORYS study (2024) provides a comprehensive assessment of the micro and macro-economic returns of social protection expenditure. In particular, the study provides a quantitative analysis relying on macro-data at country level (using econometric modelling techniques and covering the EU-27) and a quantitative analysis combining individual micro-data and microsimulation and regression models, (covering the EU-27), with an analytical approach combining the micro and macro levels of analysis for four selected countries. In line with other macroeconomic estimates that do not differentiate between different types of social protection or social investment expenditures⁶, the study firsts look at social expenditure impact on economic growth (measured as GDP growth and households' income growth) at the aggregate level (not distinguishing between social investments and protection), both in the short and in the long run, including in times of crisis. In a nutshell (see figure below), there are three main takeaways from this analysis on returns to social expenditure in the scope of the econometric analysis of the study:

1. The positive impact on growth in the short term (including during the economic crisis), confirms social expenditure as a stabiliser.
2. The positive impact on growth is found also in the long term, highlighting persistent growth-enhancing effects.
3. There are complementarities between social policies, as indicated by more significant results for the overall aggregate. This also underlines the need for more in depth analysis of specific items.

Figure 2. Macro-economic analysis: impact of social expenditure on growth⁷



Source: preliminary results of the ECORYS study.

⁶ The literature points to divergin results: some growth-inhibiting, others to growth-conducive effect. See Arjona et al (2002) and Simoes et al (2015) for surveys. As already hinted at above, a key methodological difficulty lies in the fact that causality may run both ways - social spending can impact but may itself also be influenced by economic performance. Some studies find that effects depend on countries' level of development, with less positive findings for advanced countries. E.g. Im et al (2011) find a positive effect in developing countries and a negative one in developed countries.

⁷ Results from GMM (Generalised Method of Moments) estimation technique, for EU-27 Member States for the period 2006-2019. Social expenditures include spending on sickness/health care, disability, old age, survivors, family/children, unemployment, housing and social exclusion (Sources: ESSPROS database). The economic crisis is accounted as a dummy variable applied to the years 2009-2012 – i.e. the years of the crisis, namely the economic crisis context, have a negative impact on economic growth. All the coefficients shown are significant, except from the coefficient for economic crisis on GDP growth, which is not significant.

In what follows we further investigate the empirical evidence on social investment returns. We focus first on the evidence on economic returns of social investment policies, distinguishing between social investment (in particular, human capital) and social protection. We then collect evidence on social returns, with a specific focus on investigating policy complementarity.

Human capital as engine of economic growth

As stressed above, studies explicitly focusing on components of social investment highlight pro-growth potential. The consideration of expenditure on human capital (what we called above ‘stock’) as investment (present sacrifice for better future outcomes) is fully embraced in Economics (Barro, Sala-i-Martin 2003; Gemmell, Kneller, Sanz 2016; Card, Kluve, Weber 2018). The positive impact of improvements in human capital on potential production (in terms of increased productivity, higher activity rates, greater employability, and/or lower natural unemployment rate) is solidly supported by theoretical literature encompassing various types of economic growth models, both endogenous and exogenous. Empirical studies addressing human capital contribution from different approaches support its contribution. Production functions incorporating human capital as a productive factor on equal basis with physical capital, growth accounting exercises and empirical analyses of growth models using econometric techniques support this thesis. Growth accounting shows that improvements in workers’ human capital (included in the work quality component, what we labelled above as ‘flow’) represent a significant and sustained source of growth in labor productivity. This points to the need to foster a sustained investment effort in human capital to boost productivity.

Sakamoto (2020) analyses social investment (family support, education, and ALMP spending) in 17 OECD countries in the period 1980-2008, considering its impact on three channels of growth: productivity (MFP, a measure of technological progress or efficiency improvement), physical capital investment and labour. The study finds that all these expenditures have a positive growth effect, mainly via the MFP channel. A 10-percent increase in family support and ALMP spending leads to 0.067 p.p. and 0.044 p.p. increase in MFP growth, respectively (accounting for 10.3 and 6.8 percent of total MFP growth). Moreover, education spending has a positive growth effect via all three channels analysed. As education drives economic growth by supporting technological progress or efficiency gains, the author suggests that research and development (R&D) spending should be considered in the analysis along with education expenditures due to complementarity effects. While R&D is not a social investment category, the results indeed show a positive impact of R&D spending on MFP as well – even though no interaction effects are analysed.

ALMPs in general tend to exhibit positive effects on growth. At the same time, the effects may materialize over the long term and be interdependent with other public expenditure categories, notably education and training. Reforms and investments in public employment services (PES) and active labour market policies (ALMPs) facilitate entry into the labour market as well as job matches, help reduce mismatches, and support job transitions (easing structural adjustments in the economy), thereby fostering inclusive and sustainable growth.⁸ Recent studies show that ALMPs can be cost-effective in the longer run, and some can be self-financing (Brown, Koettl, 2012). Measures that are particularly relevant in this respect include, for instance, those related to the provision of effective support to job-seekers and skills intelligence, and integrated employment and social services, as well as their digitalisation (Ekkehard, Merola, Reljic, 2022).

Other examples can be also taken in relation to education expenditure. Both reforms and investments in quality education (including affordable childcare) and training are of utmost

⁸ See also [Commission Recommendation \(EU\) 2021/402 of 4 March 2021 on an effective active support to employment following the COVID-19 crisis \(EASE\)](#) (OJ L 80 08.03.2021, p. 1).

importance in this sense. Such measures are found to *directly* foster economic growth,⁹ and increase productivity, including via stronger innovative capacity and absorption of new technologies, and also by ensuring a better matching between skills demand and supply (European Commission, 2022; Thum-Thysen, Vandeplas, 2019; European Commission, 2019).

For instance, the OECD has estimated the public net financial returns to education at different levels, i.e. the difference between the public costs and benefits associated with attaining an additional level of education. They find that the **internal rate of return¹⁰ of a public investment in upper secondary education is 9% for men and 8% for women, on average in the EU22.**^{11,12} Thanks to these substantial public returns, the investment in education strongly supports the long-term sustainability of public budgets.¹³ Taking the above-mentioned budgetary benefits and costs into account, the OECD estimate that the internal rate of return of a public investment in tertiary education is 10% for men and 7% for women in the EU22.¹⁴

Aggregate economic returns of education, associated with faster GDP growth or productivity, are more complex to estimate.¹⁵ An econometric analysis of 24 OECD countries over a 40-year period (1960-2000) shows that **countries displaying higher education outcomes in terms of skills** (measured through PISA or equivalent scores) **recorded higher annual real growth rates of GDP per capita.**¹⁶ A long-term simulation based on these estimates illustrates the potential gross gains¹⁷: if every EU Member State achieved an improvement of 25 points in its mean PISA score (e.g. what Portugal and Italy achieved in mathematics over the last decade), the GDP of the whole EU would cumulatively increase by up to 6%.¹⁸

Social protection: economic stabilization and allocative efficiency

While educational, training and activation policies materialize their micro- and macro returns in the medium to long term, the returns of traditional social protection policies are mostly investigated and evident in terms of macro-economic stabilization in the short term. They smooth the business cycle by compensating income losses due to unemployment. This anti-cyclical effect of the benefit system comes into play because expenditure tends to rise when there is a downturn in the business cycle and fall when there is an upturn.

⁹ The positive impact on economic growth is recorded despite the costs that may be entailed by some of these measures (Gemell, Kneller, Sanz, 2016) and possible delayed positive effects (Card, Kluve, Weber, 2018).

¹⁰ The Internal Rate of Return (IRR) is the implicit discount rate that equalises the flows of costs and revenues of a project over its whole lifespan. A negative IRR signals that the project is not financially viable, while a positive IRR should be compared to the returns from alternative investment opportunities.

¹¹ EU22 refers to the aggregate of EU Member States that are also OECD members (EU28 except BG, HR, CY, LT, MT, RO).

¹² The benefits taken into account in this estimate include the additional tax receipts expected by the state from an additional level of education, given that the individual successfully enters the labour market, while the cost comprises the direct costs of education as well as foregone taxes on earnings. See OECD (2016).

¹³ In this context, a similar argument applies to investments in the education and skills of migrants including refugees, which provide a significant return in facilitating their labour market integration, see European Commission (2016b).

¹⁴ See OECD (2016). The gender difference in returns may partly reflect the gender employment gap, with less time spent in gainful employment by women implying less time to reap the benefits of their investment in education, as well as wage differences.

¹⁵ Indeed the results of earlier studies aiming to quantify the impact of human capital on growth were mixed.

¹⁶ See Hanushek and Woessmann's (2011).

¹⁷ The gains are computed on a gross basis, i.e. without netting out the cost associated with bringing about the simulated change in education outcomes.

¹⁸ The very long time span is justified by the time needed to reform the education systems to improve education outcomes, the time spent acquiring skills, the working life and then the retirement period.

Impact resilience, i.e., the country capacity to absorb systemic shock, is closely associated with high government expenditures on social protection. Both structural and emergency social transfers stabilized Gross Household Disposable Income during the depths of the financial and the COVID-crises (ESDE 2014, 2022) when the latter is accompanied by inclusive schemes which also secure the most vulnerable groups (Alessi Benczur Campogno Cariboni Manca Menyhert and Pagano 2022). After the pandemic the stabilization function also had a significant and positive impact on growth via the income and consumption channel: by guaranteeing income during unemployment, illness, etc., the benefit system stimulates demand and restores economic growth.

While the literature focuses mainly on the anticyclical effects of unemployment benefits, other studies indicate that social expenditures can have a stabilisation function also in other domains. In the analysis of 21 OECD countries data for years 1982-2003, Darby and Melitz (2008), using FE model with system GMM, find a stabilising function of social expenditure in pensions, healthcare, sickness and disability areas along with unemployment related spending. Those expenditures together are estimated to compensate 0.3 p.p. for every 1 p.p. of the output gap, with retirement spending alone accounting for 0.15 p.p. However, these effects vary across countries. Positive social protection expenditure's impact on economic growth, mainly attributed to health, unemployment and survivors benefits, is also highlighted in Furceri and Zdzienicka (2012) (see box 1 below).

Box 1. The Effects of Social Spending on Economic Activity: Empirical Evidence from a Panel of OECD Countries (Furceri, Zdzienicka, 2012)

MAIN PROBLEM AND RESULTS

Over the past two decades, a substantial amount of empirical research has been directed towards estimating the effect of fiscal policy on output and the existence of crowding-out (negative) versus crowding-in (positive) effects on private consumption and investment. While most of this research has focused on aggregate spending, several empirical studies have tried to identify the components of public expenditure that have the largest effects on economic activity. Among these components of public expenditure, social spending has been almost ignored by this stream of literature or it has analysed social spending solely in the context of income inequality and poverty reduction. In contrast, Furceri and Zdzienicka (2012) look at the short-term impact of social spending on economic activity in nine policy areas: old age, survivors, incapacity related, health, family, active labour market program, unemployment benefits, housing, and other policy areas. They find that an increase of 1 per cent in social spending increases GDP by about 0.1 percent after one year, which, given the share of social spending in GDP, corresponds to a multiplier of about 0.6. Spending devoted to health, unemployment or survivors has a statistically significant effect on economic activity.

METHOD

A two-step approach is used by the authors. In the first step, they construct a measure of exogenous social spending shock, by modelling for each country a policy rule in which the growth rate of social spending was regressed against its lagged values, the current and past growth rate of GDP, the initial level of debt, the share of social spending in GDP and election dummies. The residuals of this regression are then considered as exogenous social spending shocks. In the second step, they assess the effect of these shocks on economic activity. Data are taken from the OECD databases for the period 1980 to 2005.

CROWDING-OUT VERSUS CROWDING-IN EFFECTS ON CONSUMPTION AND INVESTMENT

The results suggest that while social spending has a positive and significant impact on consumption (mostly determined by health and unemployment benefits spending), it has a negligible and insignificant impact on private investment (with spending devoted to health being the only category with a significant and positive effect). These results are consistent with the idea that unemployment benefits are measures targeted to liquidity-constrained agents, for which the increase in consumption is expected to be the largest. The fact that total social spending does not have a statistically significant effect seems to be driven by the large negative effect of old age, which accounts for a large part of total spending.

KEY TAKE-AWAY

The results suggest that social spending has a significant short-term effect on output. This effect is larger in periods of severe downturns, while it is similar between countries with low and high debt-to-GDP ratios, and between countries with large and small levels of trade openness. However, while social spending has a positive and significant impact on consumption, it has negligible and insignificant effects on private investment.

There is a long standing and strong concern that the benefit system has a disincentive effect on labour market participation and therefore on the employment rate (e.g. OECD 2004). This would have negative impacts on the growth potential as well as on government expenditure levels. However, also on the micro level impacts are often not conclusive, multifaceted and next to potential negative economic impacts there are also important potential positive impacts. Sinn's theoretical model indeed shows that the welfare state fulfils a risk reducing function (Sinn, 1995, 1996). When protected by the benefit system, people engage in risky and profitable economic activities which they would probably not undertake otherwise. Redistributive social benefits, or family and children support, may also indirectly affect growth through the consumption channel, but also indirectly via productivity channel in the long term by promoting a more equal access to education – an important driver of economic growth in developed economies. Similarly, while unemployment benefits may have adverse impacts on short term labour market transitions, they can also support access to training and better matching jobs that lead to productivity gains in employment and bring positive effects in the long term.

Many of the differences observed in the estimates from the literature can also be arguably attributed to a **different problem definition and model specification across studies**, as well as **structural differences between countries**. For example, the level of decentralisation of public expenditure in social protection has also been shown to positively correlate with economic growth (Cingano, 2014).

Social expenditure may also affect economic growth indirectly by **reducing inequalities**. Cingano (2014) estimates a growth model, derived from an augmented Solow model, where growth is a function of initial inequality, income, human and physical capital. The analysis of the effect of income inequalities on economic growth on panel dataset of 31 OECD countries over the period 1970-2010 confirms a negative relationship. A particularly strong effect is observed when the income gap between the low-income households and the rest of the population is high. Cammeraat (2020)¹⁹ finds no significant effect of social protection expenditure on economic growth, but significant effects on reducing inequality and poverty. At the same time, other studies show negative effects of redistribution on economic growth (Sakamoto 2020).

Inequality in the distribution of income and wealth may indeed affect economic growth (see also Box 2 below).²⁰ Many recent studies point to a growth-reducing effect of inequality (see OECD (2015)).²¹ This also suggests that policies that merely pursue distributional objectives can also support economic performance and economic growth.²² While this is a long-standing debate²³ a number of studies support the view that inequality hampers growth over the longer term. Yet a positive impact was also found in a number of studies, especially over shorter horizons.²⁴

¹⁹ Cammeraat, E. (2020) The relationship between different social expenditure schemes and poverty, inequality and economic growth. *International Social Security Review*, 73: 101– 123.

²⁰ See Ostry et al (2014) and Ehrhart (2009) and the references they provide.

²¹ See Halter et al. (2013); Cingano (2014); Ostry et al. (2014); Dabla-Norris et al. (2015), OECD (2015).

²² For instance, "redistribution appears generally benign in terms of its impact on growth; only in extreme cases is there some evidence that it may have direct negative effects on growth" (Ostry et al. (2014), p. 4, see also OECD (2015)).

²³ See e.g Cingano (2014) and Neves and Silva (2014).

²⁴ See Halter et al (2014).

For instance, the literature refers to a possible positive effect of inequality on growth through adverse incentives (for innovation and entrepreneurship) as well as through saving and investment (for instance as well-off people save a higher fraction of their income). It has also been argued that inequality can deter physical and human capital accumulation in the presence of imperfect capital markets. Further, politico-economic channels have been highlighted, including the argument that inequality could reduce investment through higher redistributive tax pressures.

Box 2. Inequality reduction as economic growth factor (Cingano 2014)

MAIN PROBLEM AND RESULTS

As shown by Cingano (2014), rising inequalities can have detrimental effects on the long-term economic growth. The estimated effects of reduction of disposable income inequality over the period of 40 years in 31 OECD countries (years 1970-2010) amount to 0.15 percentage points increase in economic growth per year for every 1 point reduction of Gini coefficient – potentially resulting in cumulative growth of 0.8 p.p. in the following 5 years.

METHOD

The analysis is based on a growth model, derived from an augmented Solow model, where growth is a function of initial inequality, income, human and physical capital. The estimates are calculated using fixed-effects panel regression with system GMM.

INEQUALITY, SOCIAL MOBILITY AND GROWTH NEXUS

Potential effects of income inequality reduction on economic growth are strongly related with the concept of equality of opportunity. This is highlighted by further findings. Reducing inequality at the bottom of the income distribution has a relatively greater positive impact on economic growth: reduction by half a standard deviation is linked to average growth increase by 0.3 p.p. annually (cumulated increase of 7 percent over 25 years). Moreover, no negative effect of level of redistribution on economic growth is found.

The study demonstrated the negative effect of inequality on economic growth through the human capital channel: increased inequalities affect the educational outcomes in the poorer households but not in the richer ones. Estimated inequality increase of around 6 Gini points lowers the chances of obtaining tertiary education by 4 p.p. for individuals with parents of low educational background (and does not affect those chances in case of medium or high background), slowing down the society's economic performance potential.

KEY TAKE-AWAY

The study highlights the need to consider inequality to a greater extent both in social protection expenditure analyses and in social policy design. Importantly, targeted social policies in the area of education or family may induce economic growth by promoting equality of opportunity and facilitating intergenerational human capital accumulation in the poorer households.

A 2019 OECD report on the relation between social protection and inclusive growth summarizes the different possible relations. Although drafted from a development context, the conclusions overlap to a large extent with the literature outcomes for the more developed countries.

Box 3. OECD (2019). Can social protection be an engine for inclusive growth?

The transmission channels may operate at the micro, meso and macro levels. One way social protection can influence inclusive growth is through its direct impact on individuals and households. At such individual and household (micro) level, a pure growth effect may be expected by: (i) enabling households to accumulate productive assets, (ii) preventing the loss of productive capital after a shock; (iii) enabling innovation and entrepreneurship, (iv) affecting labour market participation and savings and (v) supporting investments in human capital. While most of these factors are expected to have a positive impact on growth, the positive growth effect may be moderated by a possible negative growth effect of social protection induced by a decline in labour force allocation and savings, creating dependency and adverse incentives to work and save.

Such a growth effect induced by social protection may further interact with an effect on inequality. Social protection,

especially social assistance, can indeed contribute to make the positive growth effect equalising through two main complementary paths. First, by guaranteeing a minimum level of economic and social wellbeing, serving not only as safety nets for low- income and vulnerable households and individuals to mitigate the risk of poverty, but also as spring boards that enable social mobility and help close inequality gaps. Second, by enabling equal access to opportunities, thus overcoming the savings and credit constraints among less wealthy households that can prevent human capital investments and the disruption of the cycle of inter-generational poverty.

Besides the more direct effect of social protection on inclusive growth that operates at the micro level, social protection might also affect growth and inequality outcomes at community (meso) and national (macro) levels. As regards the growth effect, at meso level, social protection can generate multiplier growth effects from increased local consumption and production and enable the accumulation of productive community assets. At macro level, social protection can have significant and broad growth enhancing effects on the economy by increasing aggregate household productivity, stimulating aggregate demand and thus increasing employment, in particular through counter-cyclical spending during economic downturns, and raising consumption and income tax revenues. In addition, indirect effects such as facilitating economic reforms, building human capital, enhancing social cohesion and influencing fertility can further help spur growth.

Combining micro and macro analysis under different scenarios

The above mentioned

the potential impact of specific policies combining microeconomic and macroeconomic models, to allow to account not only for returns at the microeconomic level but also of their overall (general equilibrium) macroeconomic impact (accounting for the overall impact of the financing or on consumption). The results of a microeconomic analysis of the impact of increases in social assistance and unemployment benefits, combined with expenditure on active labour market policies (ALMPs) and childcare were aggregated to the economy level, based on Computable General Equilibrium (CGE) estimates for selected country case studies

Social investment employment, poverty, inequality and health returns

While the former section looked into the impact of social investment on overall economic growth, this section reviews estimates of the returns in terms of the impact on main socio economic dimensions such as employment, inequality and poverty reduction, as well as other aspect aspects such as health outcomes

Employment effects of complementary social investment policies

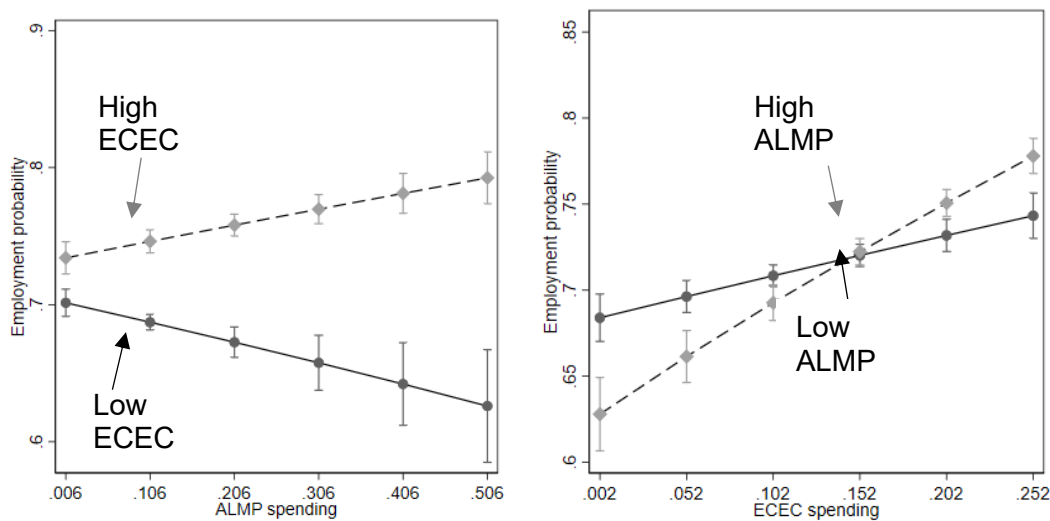
Assessing the impacts of social protection and/or social investment on employment is not straightforward as the impact of the various policies on employment transits through different channels, impacting in particular on the labour market transitions of beneficiaries (such as probability to find a job), but also more generally on labour supply (positively for instance with childcare or ALMPs) and on labour demand (depending also on how the various policies are financed). The assessments are thus often focusing more on specific areas, such as typically r ALMPs (see Martins and Costa (2014), Mörk et al. (2021) Petráš (2019)), but also other areas such as various social services including most prominently childcare.

Following the life course multiplier logic, however, we can look also at the combination of social investment policies on and how they affect people's employment chances. Hemerijck et al. (2024) explore how national active labour market policy (ALMP) and early childhood education and care (ECEC) spending efforts jointly affect employment chances among individuals with children in Europe. The authors' analyses show that in European societies, national ALMP and ECEC spending efforts promote individuals' employment chances in reinforcing ways. This is captured by marginal effect estimates plotted in Figure 2. Both ECEC and ALMP efforts are associated with a higher individual employment likelihood, but the marginal effect of higher spending effort in one

policy increases at higher spending levels in the other policy. This pattern is stable across all subsamples and alternative model specifications.

The complementary policy effects are somewhat stronger among female respondents, and especially strong for how ECEC efforts moderate (strengthen) the generally positive ALMP effect on employment probability. To visualise this, Figure 2 graphically plots predicted probabilities to be employed among female respondents at different levels of national ALMP and ECEC spending levels. In the left panel, we plot probabilities across the full sample distribution in ALMP spending in settings with low and high levels of ECEC spending efforts. In the right panel, we plot the same across the full distribution in ECEC spending in settings with low and high levels of ALMP spending efforts.

Figure 3. Predicted probability of employment among female respondents with children for ALMP effects by ECEC spending level (left-hand panel) and for ECEC effects by ALMP spending level (right-hand panel)



The plotted estimates in the left panel show a clear pattern of the complementary role of ECEC on the association between ALMP spending and employment. In contexts with high ECEC spending (the 90th percentile), employment probability among women with children is high and increases with higher ALMP spending levels, from 74 to 79 per cent across the ALMP spending distribution (from 81 to 87 per cent for the total workforce including male respondents, see Figure A4.1). By contrast, when ECEC spending effort is low (10th percentile), employment probability is overall lower and decreases with higher ALMP spending levels. Such patterns suggest that at low ECEC levels, ALMP might be unemployment- or inactivity- promoting for families with children, especially among mothers, whereby ALMP-enhanced employment take-up is likely to be concentrated among adults without care responsibilities.

The right-hand-side panel graphically shows a similar complementarity pattern in how ALMP enhances ECEC's employment promotion among female respondents. We see that ECEC is significantly more employment-promoting at higher levels of ALMP effort. When ALMP effort is high (the steeper dashed line), the positive association between ECEC spending and employment is significantly stronger than when ALMP effort is low (the flatter line). In settings with high ALMP spending, the predicted employment probability among women with children increases by fifteen percentage points across the ECEC spending distribution (from around 63 to 78 per cent). In settings with low ALMP spending, the estimated increase is only around six percentage points. The complementarity dynamics are weaker but still significant for the total workforce, pointing at

somewhat weaker ALMP–ECEC policy interdependencies for male respondents with children (see Appendix, Figure A4.1).

We can conclude from these empirical analyses that, while national ALMP efforts are generally positively associated with increased employment chances, the likelihood of being or choosing to be employed tends to significantly increase by public ECEC efforts that allow combining employment activity with family duties. Likewise, parents of children in contexts with higher national ECEC efforts are more likely to find and keep a job, but this probability is substantively higher at more elevated levels of national ALMP efforts to up-skill and ease transition into employment. This reinforcing policy effect shows up for all working-age individuals with children, particularly among women.

Poverty reduction effects of complementary social investment policies

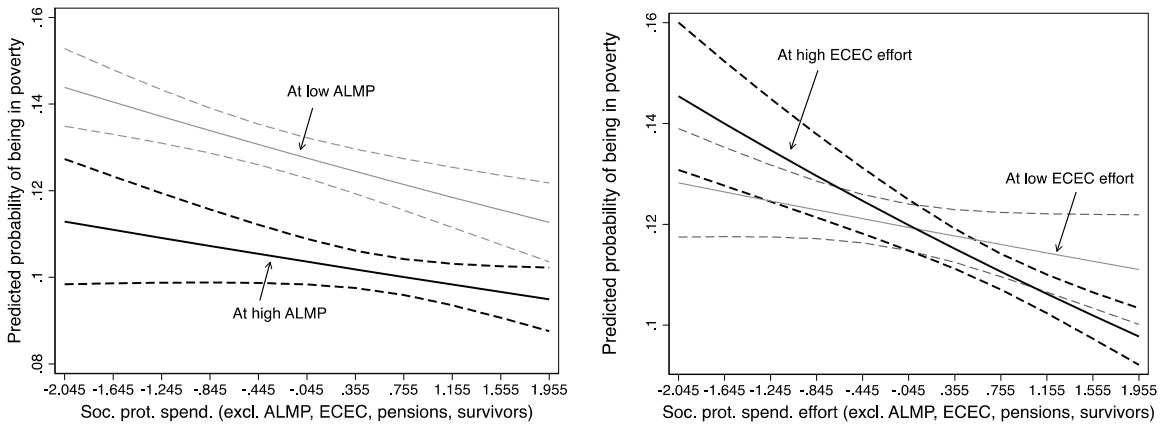
In general, compensatory social expenditure is associated to **income redistribution** and **reduction of income inequality**. For example, Niehues (2010), in a study of 24 European countries in the period 1993-2007, finds that social welfare spending, especially in unemployment and pension areas, has a redistributive effect. Nicaise and Lehweß-Litzmann (2018) focusing on the period 2006-2015 highlight that on average, the effectiveness of social benefits in alleviating poverty has been eroding over time, while the poverty gap is increasing in all but 6 countries. These averages hide an heterogeneity at country level: while there is deterioration for 13 countries, for another 13 there is an improvement. Similarly, Cantillon et al. (2012) demonstrate the poverty reduction capacity of social security expenditures across the EU over the period of 1995-2008 but identify different trends across several EU country clusters. Caminada and Goudswaard (2009) analyse the comparative effectiveness of anti-poverty social policies, including transfers and taxes, in EU15 and 11 other OECD countries, and find that the effect of transfer policies among different low-income groups shows significant variation across countries. Diris and al (2017) point to a similar conclusion.

While the above findings relate to traditional compensatory policies and their effectiveness, some observers worry that social investment can have “Matthew effects” where social investments hurt the most vulnerable citizens, traditionally most dependent upon welfare state provisions (Bonoli et al., 2017; Bonoli and Liechti, 2018; Cantillon and Van Lancker, 2013). Such effects might arise not only by potentially crowding-out spending on social-protection (buffer-focused) policies, but also through possible conflicts – actual incompatibilities – between social investment and social protection: Deepening of social investment policies like ALMP and ECEC might tend to diminish the poverty-fighting efficacy of social protection buffers like unemployment and family insurance transfers. There are good reasons to be doubtful of their severity and to expect instead patterns of complementarity between the poverty-fighting effects of social investment and of social protection. First, both ECEC and ALMP as social investment interventions focused on labour market activation and work-family reconciliation that serve as spurs to flows and stocks may well also have short-term positive effects on earned income – and hence operate as poverty-fighting. Second, even if the aforementioned incompatibilities occur empirically, the interactions between social investment and social protection in shaping poverty risk may well be accompanied by important forces of complementarity. To name but one example, vulnerable unemployed citizens able to make use of ALMP and ECEC social investment provisions can be expected at the margin to more quickly leverage social protection assistance to seek and reach income gains in labour markets.

Hemerijck et al. (2024) show indeed that well designed social investment policies are not incompatible with poverty fighting performance. The figure below focuses on the low-educated sub-sample – a key economically vulnerable group of concern among those fearing ‘Matthew

effects'. The two panels show how the full sample variation in social protection spending effort predicts lower probability of a given respondent being in poverty in settings with low (bottom 10th percentile) and with high levels (top 90th percentile) of social investment with respect to ALMP (left-hand panel) and ECEC (right-hand panel).

Figure 4: Conditional effects of Social Protection on probability of poverty among low-educated, in high vs. low ALMP settings (left-hand panel) and ECEC settings (right-hand panel)



While the two panels capture quite different joint and interactive effects for poverty-fighting, they both show complementarity in their joint effects. We can see in the various schedules that settings characterized by generous social protection *combined with* generous ALMP or ECEC (the darker schedules in the panels) are associated with lowest probability of an individual falling into household poverty compared to individuals in settings marked by social protection without generous social investment. In fact, our analysis also shows that social investment measures of ALMP and ECEC show signs of poverty-fighting in their negative associations with individual-level poverty for all of the sample populations – and in across many subsamples are revealed to be more consistently poverty-fighting in the current data and period of analysis than the most encompassing social protection buffer measure.

More important are the two broad results for how ALMP and ECEC might moderate the poverty-fighting efficacy of social protection buffers. There we see contrasting results for the ways ALMP as opposed to ECEC moderate the poverty-fighting effects of social protection. Hemerijck et al (2024) find that ALMP tends to have no statistical or substantive moderating effect on how social protection measures affect, or are associated with, poverty risk – that is, betraying neither complementarity nor incompatibility in how ALMP-based social investment moderates poverty-fighting effects of social protection. This can be seen in the left-hand panel of Figure 3, in how social protection’s negative association with poverty is only very slightly stronger in the low-ALMP settings than the high-ALMP settings. Even in the latter, generous ALMP settings we see (marginally) significant negative effects of social protection on poverty risk. This pattern holds across all of the samples in our estimates – even, for instance, among the unemployed population that might be most victim or subject to the Matthew effects logics sketched above.

On the other hand, Hemerijck et al (2024) do find a significant moderating effect of ECEC effort on the poverty-fighting function of social protection – but in the direction of poverty-fighting complementarity rather than Matthew-effects incompatibility. The right-hand panel of Figure 1 captures this pattern: there we see that social protection’s negative association with poverty risk is significantly more negative in settings characterized by generous ECEC than in settings with meagre ECEC spending effort. This is a clear and strong moderation pattern of complementarity. This pattern holds across most of the subsamples: for all but the unemployed sub-sample, social

protection has a statistically-significantly more negative association with poverty probability in settings with higher than with lower ECEC effort.

Health outcomes

Another stream of research focuses on health outcomes that can also be seen as a proxy of society's development and wellbeing. The dimensions considered most often relate to life expectancy and mortality rates, and at times also more complex ones accounting for various elements of the health status.

For instance, **cost-effective public provision of healthcare enables working-age people to remain active and at their full capacity for longer.**²⁵ OECD²⁶ estimates that about 555.000 people aged 25-64 (the working age population) died from major non-communicable diseases²⁷ (NCDs) in the EU in 2013. Assuming that these people would have been employed until 65, the associated potential loss for the economy is estimated to be around 5.2 million potentially productive life years in 2013 across the EU28 countries.

As regards **health (public) expenditure** Leitner and Stehrer (2016), analyse for EU countries for 1995-2014 links between public health outcomes and health expenditures. Health spending is positively correlated with life expectancy (and negatively with mortality rates). Some substitution effects are identified between public and private expenditure: households spend more on healthcare when public expenditure is lower. Another study on 14 OECD countries over the period 1996-2006, from Heijink et al. (2013) find that increased level of spending affects health outcomes in that it reduces the avoidable mortality from a number of diseases, where a rise in healthcare spending of 1% is associated with avoidable mortality decrease of 0.5%. They also compare cost-effectiveness of increased healthcare spending, defined as the ratio of the average gain in life expectancy and the average growth in healthcare spending. The results differ across countries, with estimated costs between USD 10,000-50,000 (PPP) per life year gained. As the heterogeneity between the countries is very high, and their analysis covers only certain diseases across rich countries with fairly similar healthcare systems, the results for a different set of countries could differ. Linden and Ray (2017) consider a larger sample of countries over a longer period, using data from 34 OECD countries over four decades (1970-2012). The findings highlight that in countries with high public health expenditure, both public and private expenditure affect life expectancy positively. However, in countries with lower public health expenditure, private spending has no additional effect – or even negative effect in the countries that spend the least on public healthcare. Private healthcare expenditure has a positive effect on life expectancy only in countries with public healthcare expenditure above around 7.5% of GDP, and the correlation is negative in countries with healthcare spending as share of GDP below around 5%. The drivers of these different effects is however not analysed.

Other studies analyse **the effect of other social investment categories on health outcomes.** In a study of 14 European countries over 1995-2017, Martín Cervantes et al. (2020), look into the relative importance of different public expenditure categories in explaining health outcomes. The most significant determinant is general social expenditure as a share of GDP, followed by health expenditure and then national income, educational levels as well as public expenditure on

²⁵ Suhrcke et al (2011). At the same time, investing in the public provision of healthcare helps limit future costs escalation related to the treatment of preventable diseases. Investments in public health and disease prevention/health promotion have a particularly high ROI as they have the potential to avoid the occurrence of premature deaths and chronic diseases, implying a loss of potentially productive life years.

²⁶ OECD and EU (2016).

²⁷ Cardiovascular diseases, cancer, respiratory diseases and diabetes.

environmental protection. Other studies show similar relations in the EU and OECD countries. Van den Heuvel and Olaroiu (2017) in a cross-sectional study, based on data from 31 European countries, identify social protection expenditure as the most relevant determinant of life expectancy, followed by expenditure on education and healthcare – controlling also for healthcare system quality and lifestyle. Bradley et al. (2011) find similar results for five health indicators (life expectancy at birth, infant mortality, low birth weight, maternal mortality and potential years of life lost), using data from 30 OECD countries over the period 1995-2005 – both health and social services expenditures are statistically significant determinants for 2 or 3 health indicators, respectively. McCullough and Leider (2016) estimate the effect of investment in healthcare as well as social services – considered broadly, including community development and culture spending – on health outcomes in the US context, using local level data from the counties over the period 1997-2015. Both social services expenditure and health expenditure have a clear positive effect on health indicators. Reynolds and Avedano (2018) look into effect of social expenditures across six domains other than healthcare on life expectancy for 20 OECD countries over 1980-2010. The impact of non-health public social investment expenditure (including education and social protection spending) is significant and is relatively greater on health outcomes than healthcare spending only.

In this respect, as far as health outcomes are concerned, overall social expenditure as well as the composition of the public budgets matter. There are important relations between different social investment categories including social services, as well as with other public spending domains (including protection of environment). One of the channels whereby other social investment impacts effects on health outcomes is by supporting (financial) access to healthcare. Israel (2016), in a cross-sectional multilevel analysis, using 2012 data on EU and EEA countries, highlights that unmet medical needs are income-related and indebted individuals are more likely to avoid seeking medical support. Overall, financial barriers to healthcare accounts for more than 50% of access barriers, facing c. 3% of the European population (Estonia, Latvia, Poland and Romania are the EU countries with the highest share of people with self-reported unmet needs for medical examination and treatment). Other determinants include also factors such as sex, household composition (single parents more likely to give up medical consultation), education, migrant status or unemployment status. Access to healthcare affects health of the society in the long-term and can also drive higher long-term care needs in older age, and more targeted income-related policies in the healthcare sector, e.g., providing free medical care for the low-income groups, can support improving overall health outcomes.

As concerns long-term care, evidence points to examples of analysis of cost-effectiveness of interventions in areas such as fall-prevention²⁸, adapted living environments²⁹, rehabilitation, healthy nutrition and obesity prevention³⁰, prevention through early screening³¹, as well as adult learning activities for the elderly. The role of digital technologies in the health and long-term care sectors is considered promising and could serve as an important driver to increase the cost-effectiveness and quality of certain long-term care interventions or measures.

²⁸ Hewitt, J., Saing, S., Goodall, S., Henwood, T., Clemson, L., & Refshauge, K., ‘An economic evaluation of the SUNBEAM program: a falls-prevention randomized controlled trial in residential aged care’, *Clinical Rehabilitation*, 33(3), 2019, pp. 524-534.

²⁹ Pega, F., Kvizhinadz, G., Blakely, T., Atkinson, J. & Wilson, N., ‘Home safety assessment to reduce injurious falls in community-dwelling older adults: Cost-utility and equity analysis’, *Injury Prevention*, 22(6), 2016.

³⁰ Lee, Y., Mozaffarian, D., Sy, S., Huang, Y., Liu, J., Wilde, P.E., Abrahams-Gessel, S., Jardim, T., Gaziano, T., & Micha, R., ‘Cost-effectiveness of financial incentives for improving diet and health through Medicare and Medicaid: A microsimulation study’, *PLOS Medicine*, 2019

³¹ Thaison Tong, Praveen Thokala, Brian McMillan, Rob Ghosh, John Brazier, Cost effectiveness of using cognitive screening tests for detecting dementia and mild cognitive impairment in primary care, 2019

Some of the other links between different categories of social expenditure and health outcomes may likely be explained by long-term effects that are not directly captured by the analyses. For example, childcare has been shown to have positive health effects in children, in a paper by Van den Berg and Siflinger (2020). Expenditure on education may lead both to greater health awareness and affect lifestyles, leading to better health outcomes (Van den Heuvel & Olariu, 2017).

Key methodologies to monitor and evaluate social investments.

As observed above, social investment returns can be defined in many ways, depending on the type of policies, their economic and social outcomes, and their determinants. The results of the existing research depend on the type of available data, the scope of the analysis, and research problem at stake. This is reflected in a wide variety of analytical approaches.

In what follows we provide a **general approach** and an **overview of different methodologies to estimate quantitative social investment returns**. It is worth noting from the start that as it is not possible to account for the returns exclusively in monetary terms, the literature relates both to the estimation of economic returns of social investment (i.e. those that for which monetisation is possible), as well as more generally returns that are not necessarily directly of a monetary nature.

While not all returns are economic or possible to grasp in monetary terms, it remains useful to underline that social investment policies also have direct economic impacts and that in particular economic value through different channels that have been identified in the literature, such as insurance and stabilisation, and the tackling of market failures (see for instance Atkinson 1995 or Barr 2012 for a review). In this respect, Atkinson (1999) underlines that a model based on assumptions of a first best economy systematically rules out the market failures that is one of the fundamental tasks of the welfare state to address : *[This] theoretical framework incorporates none of the contingencies for which the welfare state exists.*

Monitoring arrangements, indicators and the importance of data infrastructure

Before looking into the general approach and the specific methodologies for social investment returns' measurement, a short introduction to monitoring and evaluation systems can be useful.

Based on the EU Better Regulation Guidelines, **monitoring** is a continuous and organised process of systematic data collection (or access) throughout the life-cycle of any initiative to oversee its progress. Monitoring as such looks at the 'what': what has occurred; what has been the output of an intervention? Building on this, a policy evaluation assesses the extent to which the intervention is effective, efficient, relevant, coherent and brings added value compared to prospective alternatives. As put it in the Better Regulation Guidelines, *monitoring is necessary and can be used for a variety of purposes. Some of the most frequent reasons for monitoring include:*

- *Understanding of how initiatives are proceeding; it can serve as an early warning system, for instance in case of weak progress of an initiative or unexpected or unintended findings, when it may suggest a need for subsequent corrective adjustments;*
- *Evidence-informed policymaking, by tracking and reporting on implementation progress against objectives and targets and enhancing accountability;*
- *Addressing external requirements for compliance and control, such as stakeholders' demands for reporting and accountability;*
- *Effective communication to wider stakeholders, ensuring transparency on EU initiatives and helping to explain the progress towards policy goals.*

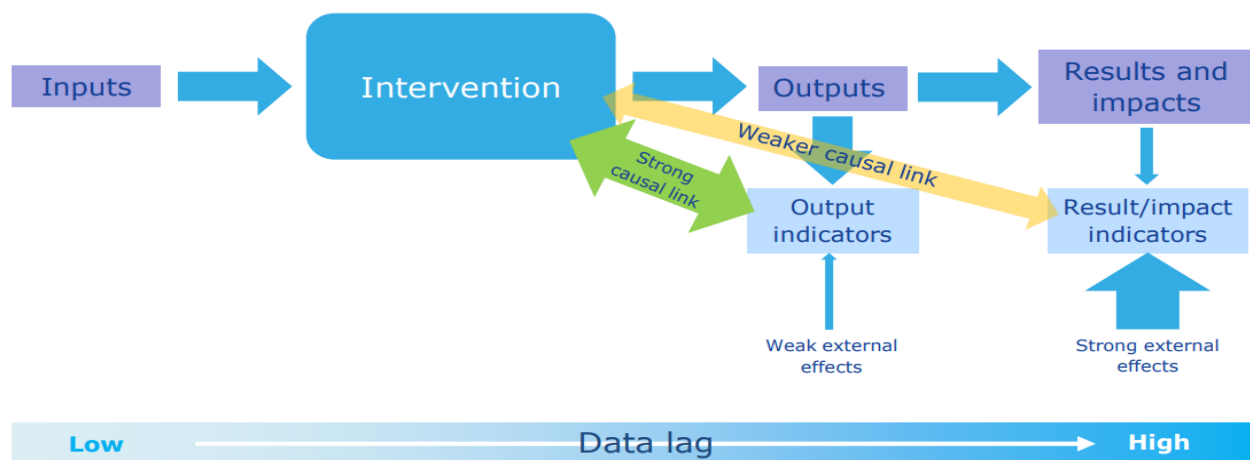
As such, monitoring is an integral part of the evaluation process. **Evaluations** go beyond a factual assessment of the 'what' to consider 'why' it has happened, how much change can be attributed to the intervention and to what extent this change meets the original expectations or projections. Evaluations aim in principle to draw conclusions about the causal effects of the EU intervention on the actual outcomes/results. Monitoring and evaluation are thus complementary, as they

ultimately aim to check whether policy objectives are being achieved. **Indicators** are the main instruments for monitoring, since they track measurable progress towards defined objectives. As such they should be clearly defined and closely linked to the activity. The measurement unit selected should be defined precisely, and indicators have to be measured sufficiently frequently to monitor progress. Four types of indicators to track spending can be distinguished:

- **Input indicators** traditionally relate to the expenditure planned and disbursed, and thus the take-up of resources;
- **Output indicators** measure the immediate deliverables of the intervention, produced as a consequence of the initiative. They relate to what is directly produced and supplied through the implementation of financially supported operations;
- **Result** (also known as outcome) **indicators** measure the effects of the interventions supported, with particular reference to the direct addressees, population targeted or users;
- **Impact indicators** refer both to the long-term impact on participants or broader target groups in society, beyond the participants or entities directly benefitting from support.

The more we move towards impact indicators, the more difficult it is to link the initiatives to what they deliver as results and impacts. This is due to a variety of factors, making it sometimes difficult to differentiate between correlation, causality and incidental effects. Moreover, data lags tend to be longer for impact indicators. For this reason, a proper monitoring system should also collect contextual information to account for developments that are not intentionally related to the individual initiative, but can influence it. From the perspective of policymakers, indicators capturing results and impacts are the most important to monitor, yet it can be difficult to find a causal link between the intervention and its impact. In addition, depending on the initiative, it can take some time before it has an impact and even more to be able to measure the impact with an indicator. **Causality** and **timeliness** are thus two aspects that complicate any monitoring and evaluation exercise. In addition, other relevant factors that should be considered include external factors (unexpected and external input that influences the output and/or the outcome) and side effects (unintended results of the used up input). The figure below illustrates the above described process (except side effects).

Figure 5. System view of an initiative



Source: Better Regulation Toolbox (European Commission, 2021a), p. 359.

Given that monitoring is a systematic process to track progress and generate information for evaluation, arrangements need to be set-up for collecting, processing, and using/reusing **data**. Setting up a monitoring system could benefit from IT support. Digitalization will lead to simplification, burden reduction and less errors only when the monitoring processes and related data flows are well considered and streamlined in advance. To this end, it is important that policy officers consider the use and the reuse of IT systems and reuse of data, whenever possible, as well as data protection aspects.

Capabilities of such an IT system could include:

- cataloguing data collection requirements (frequency of data provision, actors, etc.);
- collecting or harvesting data;
- data storing;
- data quality assurance, including (automatic) validation;
- data processing and analysis;
- database interoperability;
- data visualising, sharing and disseminating results;
- data access and discovery (for example by making available metadata or referencing your data on data.europa.eu).

Data to evaluate social investments: state-of-play and way forward

Before delving into the approaches and methodologies to track social investment returns, an birds' eye view on the state of play of data management and access can be useful. In the realm of social investment, the management and access to data plays a pivotal role in decision-making, impact assessment, and overall success. Social investment encompasses a broad range of policies, and access to appropriate data is instrumental in this mission.

The question of data is closely linked to that of evaluating public policy and social investment. The possibility, accuracy and quality of the results provided by the various assessment tools available (whether they are macro or micro, ex ante or ex post) are indeed largely dependent on the data available. Data play a crucial role in the development, update, and utilization of various monitoring and evaluation methods. Regarding the use of microsimulations models and more generally, Distributional Impact Assessment – which will be reviewed in the next sub-section, it is common to utilize surveys data as well as administrative data; each having its own advantages and disadvantages. The strategic choice between administrative data, known for precision, and survey data, which reveal certain otherwise inaccessible information, depends on the context and possibilities. A combination of both is also a viable option.

Survey data allow for the inclusion of specific or ad hoc questions and generally offer extensive coverage of households, their composition, and individual characteristics. Indicators based on data derived from surveys conducted on a sample of the population, such as the EU-SILC survey data or the Labour Force Survey data, carry a statistical margin of error that needs to be considered during interpretation. Additionally, variations over time and across different regions within the sample may not always be generalizable to the entire population. Surveys also face issues like low response rates, often have limited coverage of high-income individuals and can be quite expensive.

Administrative data, on the other hand, are relatively cost-effective as they are routinely collected by public administrations. They help reduce measurement errors, eliminate recall errors, provide comprehensive coverage of the population, and can capture the high-income distribution effectively. However, they often have a limited coverage of the bottom tail and can pose challenges in identifying families and households.

Over the last decade important steps have been made at the European level to enhance statistical capacity required for social policy impact assessments. Importantly, there is the 2019 Framework regulation on social statistics ((EU) 2019/1700), which offers a sound legal basis for EU social data collections (e.g. EU-SILC, LFS, EHIS,) and allows some flexibility regarding the information collected.

However, social investments include a wide range of policies in terms of types of policy and target population. To further extend the range of possible impact assessments more detailed data are required. To achieve a comprehensive analysis, it is imperative to employ a sizeable and representative sample that encompasses all the requisite components of the policies under examination and takes sub-populations into account. For this purpose, enhanced availability of administrative data for policy impact assessments is a priority. Currently, the use of administrative data for policy development differs strongly between member states.

While in a number of member states extensive administrative data is widely used, in many other member states there remain important barriers for their effective use. Both privacy considerations as well as institutional rigidities play a role. The resistance of some government agencies to share their data or collaborate with third parties to evaluate their public policies is a topic of ongoing debate in the field of public administration and governance. This resistance can be attributed to various reasons, including concerns about privacy, data security, bureaucracy, and a lack of incentives for transparency, competition between institutions, etc. The General Data Protection Regulation (GDPR) adds a layer of complexity to data accessibility, presenting a notable challenge, particularly in a European context. Finding a sound approach at the European Union level to address GDPR issues in administrative data, while setting standards for their use is essential to facilitate meaningful data-driven endeavors. Support for further opening up administrative data for secondary use, while fully complying to GDPR is in line with recent initiatives, launched by the Commissions' communication on an EU data strategy (EU(COM) 2020/66) and with the Data governance act ((EU) 2022/868), which became applicable in September 2023. The Data Governance Act seeks to *increase trust in data sharing, strengthen mechanisms to increase data availability and support the set-up and development of common European data spaces*. The idea of the initiative is to *make more data available and facilitate data sharing across sectors and EU countries*. Notably, work is in progress on the European Health Data Space (EHDS), designed to enhance the quality of public finances. The latter has been proposed in 2022 by the Commission (2022/0140(COD) and is still under negotiations. This space encourages national decision-makers to delve into the realm of health data, promoting the acquisition of more and higher-quality data for informed decision-making and policy formulation. The EHDS aligns with the broader vision of social investment, emphasizing the need for accurate and comprehensive data to guide informed policy decisions and improve the well-being of citizens. The Data Act, which entered into force in January 2024, further complements the data framework and aims to regulate who can use which data under which conditions. Currently a Commission proposal amending Regulation (EC) No 223/2009 on European statistics is under discussion. The proposal aims at improving the access, reuse and integration of administrative data for the development of European Statistics. The latter is of particular interest to enhance the use of administrative data for policy monitoring. However, much will depend on the scope of the amendment regarding data and the actors that will have improved access to the data.

The Communication on the EU data strategy underlines that sharing of data between public authorities is important as it can make a considerable contribution to improving policy making and public services, but also to reducing administrative burden on companies and citizens. The strategy also puts forward the idea of creating data spaces in different sectors, among which, if adopted, the European Health Data Space will be the first of its kind. Interestingly, among the proposed data spaces figure also: Common European data spaces for public administration, to

improve transparency and accountability of public spending and spending quality, fighting corruption, both at EU and national level, and to address law enforcement needs and support the effective application of EU law and enable innovative 'gov tech', 'reg tech' and 'legal tech' applications supporting practitioners as well as other services of public interest.

A Common European skills data space to reduce the skills mismatches between the education and training system on the one hand and the labour market needs on the other. Creating such data spaces around socio-economic data and implementing the provisions of the data act, i.a. regarding data intermediation services and a related regulatory framework can be crucial enhancements of the capacity to evaluate social investments and of evidence based policy making in general.

Box 4. National data platforms

Croatia is currently updating the National Registry of Households, providing a comprehensive overview of household income. This new system will serve as a basis to identify families requiring incentives and more tailored measures to achieve national policy impacts. The registry-based national data is expected to enhance transparency, quantify and monitor policy effects on households, and provide support to the most vulnerable.

Luxembourg utilizes the Luxembourg Microdata Platform on Labour and Social Protection (LMDP)³², facilitating the provision of administrative data for academia and relevant administrations. The LMDP was created collaboratively by the Ministry of Social Security (MSS) and the Ministry of Labour, Employment, and Social and Solidarity Economy (MTEESS). The LMDP aims to facilitate access to pseudonymized individual administrative data, enhancing the statistical use of administrative data for research on the labor market and social protection. It also ensures privacy compliance with the European General Data Protection Regulation (GDPR). The LMDP is intended for experts in the fields of employment and social protection, such as researchers, study officers, and analysts, working on projects with an exclusively statistical purpose.

The **Belgian** Datawarehouse "Labour Market and Social Protection"³³ was created in 1999 at the request of several social security institutions to record their administrative data. Managed by the Crossroads Bank for Social Security, this data warehouse serves scientific purposes and facilitates the construction of usable data for research. The data within it are diverse and do not form a homogeneous and integrated database, except for what is available online. The primary objective of the Labor Market and Social Protection Data Warehouse (DWH MT&PS) is to aggregate socio-economic data from Belgian social security institutions and other public entities.

Overall approach to evaluate the return on social investments

As put it in the EU better regulation guideline, '*social investment has positive consequences for economic growth and societal well-being in general. However, as in any category of economic decisions, social investments have to consider their opportunity cost*'. Economists try to answer these questions by considering the costs and benefits linked to each option; **cost-benefit analysis (CBA)** is characterised by the attempt of measuring all effects of a policy option in monetary units. The OECD points out that when applied appropriately, CBA needs to take account of the full range of costs and benefits, including financial, economic, social and environmental. Based on the EU Better Regulation, the application of cost-benefit analysis requires the following main steps: 1) Identification and monetisation of costs and benefits; 2) Selection of the relevant time horizon and social discount rate; 3) Choice of a mathematical aggregation rule, 4) Presentation of the impacts and formulating the judgement on the performance of existing public intervention or the comparison of the policy options; 5) Checking the robustness of the results;

³² [Luxembourg Microdata Platform on Labour and Social Protection - Inspection générale de la sécurité sociale // Le gouvernement luxembourgeois.](#)

³³ [Datawarehouse | Mission \(fgov.be\).](#)

and 6) Accounting for distributional and cumulative impacts. The chart below illustrates some examples of national practices and EU initiatives using CBA approach.

	cost		benefit	
	monetary	Non-monetary	monetary	Non-monetary
Fiscal				
Economic				
Social				

As already indicated above, it is not possible to express all costs and, especially, benefits in monetary terms. Indeed some of the effects of social investments do not typically have a monetary value (for example, reduction in poverty or increased social inclusion, better health outcomes, even though those also have an economic impact) or the methodologies for estimating those are not yet well established and widely used in the literature (for example in case of improved health outcomes). For instance, the non-monetary returns can be very important for the social cohesion (e.g. reduced crime rate, higher overall wellbeing, preventing social disruptions, etc.). In many instances a CBA approach is therefore replaced or complemented by multi-criteria decision analysis (MCDA) which similarly allows considering a wide range of assessment criteria, which can be included in their original units of measurement. Social multi-criteria evaluation (SMCE), which has been explicitly designed for public policy, also allows the capture of distributional consequences (e.g. in terms of stakeholder types, EU regions/countries or time) and trade-offs between dimensions (such as between some economic, social or environmental impacts, or between some families of criteria). To guide the scope of evaluation activities for social investments and reforms a multi-criteria impact matrix may be useful. The table below distinguishes three dimensions of analysis (fiscal, economic, social). Of course, further dimensions can and should – where possible - be looked into between direct and indirect impacts and between short, medium and long-term impacts (e.g. environmental impacts). Not every cell is relevant or feasible in every situation. Which impacts can be assessed and which methods can be used will depend on the policy issue at hand and on national and EU availability of statistical capacity in terms of data, tools and analytical expertise.

A complete MCDA does not only distinguish between different impact dimensions. Further specification regards the **objective**, i.e. desired direction of a policy change. For example, within the economic dimension, economic growth has to be incentivized, in the social dimension, the fairness should be maximized, in the fiscal dimension, the sustainability should be maximized. Any objective may imply a number of different criteria. Criteria are empirical indicators that associate each policy option with a variable indicating its desirability according to its expected real-world consequences. In the economic dimension, GDP is an indicator of growth maximization. **Criterion score** is an assessment of the impact according to a given criterion with reference to each single policy option. Criterion scores can be both qualitative and quantitative; uncertainty can also be included. Uncertainty can be linked to external factors (unexpected and external input that influences the output and/or the outcome) and side effects (unintended results of the used up input).

Box 5. Examples of CBA and MCDA at national level

Some EU Member States have integrated a CBA or MCDA into their social policy planning. The **Danish socio-economic investment model** (Socioøkonomisk Investerings model, SØM) is an example of an applied cost-benefit approach. The aim of the SØM is to provide “economic analysis to anyone interested in how a social policy affects the public economy in both the short and longer term and in knowing what the economic costs and benefits of particular social investments are” (Kvist, 2018). It consists of a database of effective practices and a calculation framework to determine the implementation costs. When the user selects the policy initiative to consider, and characteristics of the target persons, the model calculates direct cost and possible outcomes. These are calculated using estimates from register data on success rates for participants and impacts on public spending such as changes in demand for public transfers/benefits, tax revenues etc. The first version, published in January 2018, included adult target groups only. The model has now been updated to include vulnerable children and young people aged 6 to 17 and will soon include children aged 0 to 5.

Similarly, in the **Netherlands** the social policy planning is underpinned by a number of cost-benefit analyses conducted by the Centraal Planbureau (the Netherlands Bureau for Economic Policy Analysis (CPB)), which is one of the Netherlands’ three national policy assessment agencies. For example, one of the recent analyses considered calculation of policy options for an alternative allowance system (CPB, 2020) and its effects on budget balance, employment and inequality (Gini coefficient). Another analysis (CPB, 2020) focused on GroenLinks social security and tax system reform and its impact on the purchasing power of households and structural labour supply. Similar effects have been investigated for social policy changes proposed by the Labour Party (PvdA) (CPB, 2020).

In **Slovakia**, a CBA methodology was applied to the assessment of value for money for interventions in the field of early intervention for children with hearing impairment (Bednarik et al., 2018). The analysis included such benefits as reduced training costs, revenues and reduced costs of the state due to higher employment of parents and of people with hearing impairment, contribution to the compensation of personal assistance. On the other side, the included costs were linked to remuneration of mobile pedagogues, travel expenses, tools and toys, education and supervision of pedagogues, program management and increased training costs. In addition, a sensitivity analysis was conducted to observe the change of the CBA results depending on the change of the values of input parameters due to external environment or other uncertainties. The study concludes that every euro invested in the early intervention for hearing-impaired children brings benefits of 1.40-5.53 EUR thanks to the reduced state’s future social assistance costs or additional revenues from taxes and levies earned by higher employment of the future adult. In **Croatia**, CBA was used to estimate the cost-effectiveness of rehabilitation programs despite the difficulties related to the lack of a unified taxonomy of rehabilitation measures (Bejakovic et al., 2013). A probability model is used to consider the probabilities of employment and remaining in the world of work and shows that seemingly low probabilities of employment over a period of one year lead to relatively high and acceptable levels of employment in the long run. Depending on the observed scenario, the difference between today’s value of cumulative cost and future benefits from investing in the rehabilitation program ranges from a minimum of 170 to even a maximum of 325 thousand kuna in case of non-payment of benefits during employment, and from 131 to 273 thousand kuna in case of payment of benefits while a person with disabilities is employed.

Overview of the approaches and methods to measuring social investments returns

Although there has been important advancements in terms of impact assessment capacity over the last decades, statistical capacity to perform impact assessments still remains very uneven among member states. In practice, and at least over the median term, a flexible approach will be required for evaluating social investments, using the already extensive, and expanding, existing instruments.

Various methodologies are employed in literature³⁴ to measure the returns of social protection expenditure, ranging from micro to macro analysis, ex-ante and ex-post evaluations, and employing regression-based methods, microsimulation, cost-benefit analysis, and Social Accounting Matrix (SAM) based modelling. In this section, we will provide an overview of existing

³⁴ As shown in the recent study commissioned by DG EMPL to Ecorys: ‘Study on assessment of micro and macro-economic returns of social protection expenditure’ (November 2023). Much of what is presented in this section stems from this study.

possible approaches and tools to assess costs and returns on social investments.

Macro-level analysis indicates positive returns of social protection expenditure on GDP growth and household income, both short and long term, even during crises. While short-term effects on employment are negative, they tend to dissipate over time.

Micro-level analysis yields mixed results on the impact of social protection on employment and poverty reduction, influenced by factors such as benefit types, combinations, and target groups.

The table below highlights the main methods available to assess the returns on social investments, combining ex-ante and ex-post approaches together with micro and macro level assessments:

- As for the micro dimension, a number of studies rely on microsimulation and quasi experimental methods.
- As for the macro dimension, a large part of the literature available relates to ex post regression-based methods, as well as micro-macro modelling.

Table 1. Typology of main methods available to assess social investment returns

	Ex-ante	Ex-post
Micro-level	Microsimulation	Quasi-experimental methods (such as counterfactual simulation and randomised control methods)
Macro-level	Macro-economic modelling (including social accounting matrixes or general equilibrium models)	Ex-post evaluation (of macro variables, including through regression models and frontier analysis)
Both	Cost-Benefit Analysis, Distributional Impact Assessment (DIA)	

Source –Ecorys (2024)

Given the difficulties to identify social investment returns, the various methods are in general very intensive both in terms of the data needed to actually be able to elaborate some (granular) estimates and the complexity of the methods needed to allow the calculation of the estimates (such as through actual experiments or through various types of ex ante or ex post estimations or modelling). The table 2 below presents some of the advantages and limitations of the main methods identified. Overall, the impact of social investment is challenging to measure and quantify, as there are many elements that can also affect economic and social outcomes and some of their effects may only materialise in the long-term. The selection of the relevant approach and factors included in the modelling is bound by these limitations and challenges. While no model can account for all these elements at once, it is important that the choices of assumptions and methods are based on informed analysis.

Table 2. Advantages and limitations of various methods to measure returns to social investment and social protection

Methods	Type of data	Main advantage	Limitations	Types of social protection returns commonly analysed
Fixed Effects (FE) panel regression	Country-year data	Reduces the bias from other unobserved variables	Requires many observations – Difficult to identify causality	Economic growth Inequality reduction
FE panel regression with System Generalized Methods of Moments (GMM)	Country-year data	Allows tackling endogeneity effects and causality problems	Requires many observations – Difficult to identify causality	Economic growth Inequality reduction
Simple regressions (bivariate or multivariate)	Cross-sectional or panel data	Applicable to different types of data	Risk of biased results from unobserved variables – Difficult to identify causality	Poverty reduction Health outcomes
Multi-level regression models	Cross-sectional or panel data	Allows detecting relations at different levels of for nested data	Requires sufficient sample size at every level– Difficult to identify causality	Poverty reduction Health outcomes
Quasi-experimental methods (counterfactual simulation, regression discontinuity design)	Cross-sectional or panel data	Allows analysing causality and quasi-treatment effects of policy change by defining treatment and control groups ex post	Requires very careful specification of treatment groups or control variables to isolate the effect	Outcomes of single policy or targeted measures, for example
Microsimulations	Micro data on units (households and persons)	Simulating effects of policy interventions on units at the micro level – clearly identifies assumptions	Inherently reduces the scope to the elements modelled	Primarily income-related outcomes, poverty, social participation
Social Accounting Matrix and micro-macro approaches	All-economy data collection framework	Allows accounting for a large set of factors, reproducing entire economy within general equilibrium (CGE) modeling	Difficulties to ensure proper modelling of all aspects and detailed measurement of returns due to high level of generality	Diverse social policy interventions
RCT	Micro data on units (households and persons)	Allows analysing causality effects of policy change by randomly choosing units for the treatment and control group	Requires very good randomization of treatment groups or control variables to isolate the effect of the policy	Outcomes of single policy or targeted measures, for example

Source – Draft Ecorys study (2023).

In addition, there is analysis that focuses on the cost of non-action, such as by providing estimates of the cost of NEETS or the cost of poverty (see for instance Ozdemir and Ward 2014). Some analysis also relates to the cost of low levels in social investments that can translate into lower growth. For instance, House et al. (2017) and Heimberger (2016) underline that stronger social investment reforms could have raised productivity, labour force participation, and employability and made it easier to meet family responsibilities (see also Darvas and Tschekassin, 2015; Leschke et al., 2012, Gechert et al., 2017 or more recently Hemerijck 2023).

Ex ante micro approaches

Microsimulation involves the application of a range of modeling techniques at the individual level to simulate changes in state or behavior. This approach utilizes microdata, which consists of data related to individual units of analysis, enabling the simulation of the effects of specific policy on each unit. In concrete terms, a microsimulation model consists of a number of units with associated identifiers and attributes (e.g., employment status, age, sex) for which a set of rules is defined and applied (e.g., tax regulations, pensions). The microsimulation results in estimating the effects of the rules on the population of units (see also box 1).

In the context provided, a distinction is made between "static" and "dynamic" models. In static models, individuals' demographic states are assumed to remain fixed, whereas dynamic microsimulation models endogenously determine demographic characteristics such as employment status, marital status, household composition, etc., during the simulation (Decancq et al., 2012). In order to run, a static model needs two main elements: a detailed individual database and modules replicating the measures and legislation. Dynamic models are useful when studying a situation over the long term : they are particularly employed to support reflections and reforms of pension systems (Bessis and Cotton, 2021).

Box 6. Microsimulation and the European static microsimulation model EUROMOD

Microsimulation is one of the relevant methods to measure returns of social protection expenditure. It has been widely used in **applied research and government decision-making** in European countries.

Since the 1980s, the use of this technique has become significantly more accessible, leading to a substantial broadening of its scope. This can be attributed, on one hand, to a more widespread and simplified access to individual data (whether derived from surveys or administrative sources) and, on the other hand, to technological advancements enabling greater computational capabilities at reduced costs (Blanchet et al., 2016).

EUROMOD is a static microsimulation model with sets of rules regarding taxes and social benefits that are specific to each of the EU Member States, allowing comparisons between member states. It is a tax-benefit model used to calculate and compare effects of existing and potential taxes and benefits policies on household income (Sutherland & Figari, 2013). The EUROMOD model was initially developed in 1996 by a team of university researchers and has since evolved through various funding obtained from the European Commission. Since 2021, the management of the model is overseen by the Joint Research Centre of the European Commission (JRC), with data support from Eurostat.

The repository of papers using EUROMOD includes 291 entries published between 2017 and 2021, of which many recent papers concern the impact of welfare policies. Several papers exploit the function of EUROMOD to describe how policy elements interact (such as for instance the analysis of how the existing tax-benefit system treats households affected by disability in Ireland, Doorley and Regan 2022). Some of the papers use microsimulation to assess the effectiveness of existing policies, such as the impact of the pension system on old-age poverty in Poland (Ratajczak and Bartkowiak 2021) or the labour supply effects of marriage bonuses in tax systems across Europe (Christl, De Poli, and Ivaškaitė-Tamošiūnė 2022). In a similar vein, microsimulation has been used to assess the resilience of welfare regimes to external shocks, such as COVID-19 (Christl, De Poli, et al. 2022), or job destruction (Christl, Livanos, et al. 2022; Colombino and Islam 2022). Other papers use it to compare possible policy options, e.g. of fiscal policies to improve childhood health (Rasella et al. 2021), or the impact of variously defined minimum-wage rules on in-work poverty, wages and employment across the EU (Grunberger et al. 2021).

In another study, EUROMOD was used to analyse **financial vulnerability in the context of the effects of the COVID-19 pandemic** on earnings' reduction in seven EU countries (Midões & Seré, 2022). The model indicated that in the wake of the pandemic, certain groups were particularly likely to be financially vulnerable: migrants, single parents, single-earner households, households with low educational attainment level, and women. In all countries the number of vulnerable people significantly dropped when awarded COVID-19-related unemployment benefits. Employment protection schemes and generous layoff schemes were found to reduce the percentage of vulnerable persons, while the effect of rent and mortgage suspensions varied among the countries included in the study.

While the classic application of microsimulation is to study impacts on incomes, poverty and inequality, it can also be used to analyse returns on social investment, including e.g. through effects on increasing labour supply, savings on energy consumption or improving health conditions. Labour supply effects are not necessarily positive, especially in the short run, as cash benefits can also reduce labour supply. Tax changes and cash benefits may have positive effects that often only materialise in the medium or long run. For example, unemployment benefits may support jobseekers in finding a higher paying job that better matches their skills, or in sustaining investment in the education of their children, or in their own health condition. Welfare services typically contribute to improving the employability of recipients or by freeing people of care obligations in their family.

Member States wishing to import or adapt the model can also be funded to do so. Today, there are many examples of EUROMOD based models. We'll be taking a closer look at some of them in some of the following boxes of this paper. In Belgium, this is the case with the BELMOD project. The FPS Social Security developed the BELMOD model in cooperation with the Federal Planning Bureau and the Universities of Antwerpen, Leuven and Essex. The model runs on a very large administrative dataset via the Datawarehouse Labour Market and Social Protection.

The Belgian Federal Planning Bureau of Belgium developed several micro-simulation models. One of the most recent is the EXPEDITION model (EX-ante simulation of Policy reforms and an Evaluation of their Distributional Impact on Income and Other welfare Notions). Based on the EUROMOD technical platform, EXPEDITION is specifically designed to assess the potential effects of policy proposals put forth by political parties, operating on a substantial administrative dataset. This comprehensive approach enables a thorough evaluation of the implications of various policy changes on income and other welfare metrics. MIDAS (Microsimulation for the Development of Adequacy and Sustainability) is another microsimulation model developed at the Federal Planning Bureau. It was built, at the end of the 2000s, to assess the social sustainability of the pension system. Unlike the models mentioned above, MIDAS is a dynamic micro-simulation model that simulates life courses and tracks transitions in household formation, the labor market, and income.. The simulations take into account new births and migration, using data from both administrative records and the Census. Additionally, it integrates data from the in-house Macro-model Malthèse.

Some examples of the use of those models in Belgium: The impact of in-work benefits and their design on employment and poverty (Vandelannoote, D. & Verbist G. (2018), 'The design of in-work benefits: how to boost employment and combat poverty in Belgium?', in B. Cantillon, T. Goedemé, J. Hills (eds.), Decent Incomes for all. Improving Policies in Europe, Oxford: Oxford University Press, pp.383-416) ; Inactivity traps in the Belgian Sickness and Invalidity system (Hufkens, T., Buysse, L., Van Mechelen, N., Verbist, G., Fit for the labour market? An effort to reduce inactivity traps in the transition from benefit to work in the Belgian sickness and disability system, Working Paper, N° 17/11, University of Antwerp).

Ex-post micro evaluation

Quasi-experimental methods estimate quasi-treatment effects of policy interventions and thereby have a key advantage to allow to estimate with a relatively high degree of robustness some causality links. Building a counterfactual scenario in social analysis is however difficult, as isolating the effects of policy intervention from possible other determinants may face a number of operational barriers, unless the intervention is designed with a "treatment" group and a "control" group. One solution to isolate the effects of policy intervention is indeed the use of **counterfactual simulation**, based on the pre-intervention data, well-defined assumptions and careful model specification that includes other potential determinants of the observed outcomes.

In situations where it is not possible to conduct randomised control trials, quasi-experimental counterfactual impact can provide valuable information for making informed policy choices, establishing the accountability of public expenditure, and building support for continued deployment of measures in the longer-term. A key requirement for such evaluations is the use of rich data, in particular administrative data, of good quality, which can be obtained by linking

different registers, or through surveys. From 2020-2024, the OECD and the European Commission are undertaking a joint project on counterfactual impact evaluation through the use of linked administrative and survey data. Within the context of the project, active labour market and social policies in Lithuania, Finland, Ireland, Greece, Portugal and Slovenia will be evaluated, focusing on employment, earnings, and occupational mobility outcomes. Overall, the project has a dual aim of improving the effectiveness of active labour market and social policies based on the results of the counterfactual impact evaluations and to strengthen the countries' capacity for evidence-based policy-making, also using linked administrative data.

Assessing impacts of policies requires a careful research design to identify causality – with a careful definition of a counterfactual scenario. Studies on single programs, that collect data on a group of participants, can **include quasi-experimental designs with participant score matching to correctly attribute** the observed effects to a given program. **Regression discontinuity design** is one of these methods, analysing the change in observed outcomes in a close interval before and after the intervention with regard to the predicted pre-intervention trajectory. Another quasi-experimental design is **propensity score matching**, which observes the effects of targeted policy or program on participants as compared to similar individuals (in terms of relevant socio-economic characteristics) not subject to the intervention. The results of such studies can be very informative - however, they are highly context-dependent and therefore difficult to generalize. Some of the reviewed studies include also other econometric techniques, based on advanced computational techniques, for example multivariate adaptive regression splines (MARS) or fractional response model (FRM).

As regards the challenges, while these approaches are insightful and allow to account for causality, their results are also highly contextual and difficult to generalize as many country-specific and other structural characteristics may determine the outcomes. Another challenge in ensuring the comparability or the generalisation of those results lie in data collection – while there are standardized datasets regarding many aspects of social expenditures across the EU, this is not the case when it comes to broader social policies and local level data.

Randomized Control Trials

A Randomized Controlled Trial (RCT) is a robust method used to evaluate public policies, programs, or interventions. It is considered the gold standard in empirical research for assessing causal relationships between a policy change and its outcomes. RCTs involve randomly assigning participants or units into two or more groups: a treatment group that receives the policy or intervention, and a control group that does not.

The key components of an RCT include:

1. **Randomization:** Participants are randomly allocated to groups to ensure that the groups are comparable at the outset, reducing the risk of bias.
2. **Control Group:** The control group serves as a baseline against which the outcomes of the treatment group are compared. They do not receive the policy or intervention.
3. **Treatment Group:** This group receives the policy or intervention being studied.
4. **Outcome Measurement:** Researchers carefully measure and compare outcomes in both groups to assess the policy's impact. These outcomes can be quantitative data, preferably administrative records, but also surveys are used, or qualitative data like interviews.
5. **Statistical Analysis:** Statistical methods are employed to determine if any observed differences in outcomes between the groups are statistically significant and not due to chance.

RCTs provide robust evidence of a policy's effectiveness, allowing policymakers to make informed

decisions based on concrete data. They are particularly useful when assessing the impact of new policies, social programs, healthcare interventions, and more. However, conducting RCTs can be resource-intensive and may not always be feasible. Pilots based on RCT on social matters has proved controversial (e.g., the Finnish experiment of giving a basic income to a certain and limited group; and compare their behaviour with those who did not receive the benefits).

Box 7. Example of RCT in Spain to support to the early childhood education in families in situations of social vulnerability due to poverty

This program project proposes to experiment with and evaluate a low-cost, highly replicable support model for '0-3 Families'. The model combines ongoing social support throughout the program with the development of workshops in the so-called '0-3 Child Spaces,' aimed at enhancing parental skills. This proposal for socio-educational action, focused on improving parental competencies for children aged 0-3 and empowering parents, is justified by the insufficient coverage of resources addressing these educational needs for children aged 0-3, especially in low-income families, and the scarcity of alternative models to 0-3 schools or nurseries

The program has a detailed memory of the interventions to be carried out, which includes:

- A. **Social accompaniment**, that include the following tools: a.- Social assessments; b.- Elaboration of a work plan including services such as psychotherapeutic support, training ...; c.- Follow-up interviews; d.- Evaluation.
- B. **Family workshops.**

There is a list of services or incentives to facilitate participation and to avoid the dropout from the project. A comprehensive description of the recruitment process of participating families is also included.

Investing amount: The pilot project of PolicyLab had a cost of 2.7 million euros, covering a total of 1,009 households that participated in the project, divided into 20 groups in different locations. The goal of this new investment is to scale the experience nationwide, and it is estimated that it will reach a total of ~100,000 interventions over the course of three years. The planned investment in this program is 300 million euros.

The results of the evaluation carried out in the Policy Lab will contribute to the definition of the corresponding metrics (outputs and outcomes), taking into account the criteria outlined in the Annex 7 of the Proposal for a Regulation on the effective coordination of economic policies and multilateral budgetary surveillance.

Regarding data collection it has been designed a map of variables with four categories: identification, participant characteristics, intervention and indicators. Information pertaining to the first two categories will be generated during the recruitment phase and will encompass records acquired through surveys, interviews, and administrative records. These latter ones are obtained after entering into an agreement with Social Security, and it is not necessary to obtain data from the Treasury, as the group is already covered as beneficiaries of the Minimum Income Scheme administered by Social Security. Information exchange procedures are enabled taking into consideration the presence of specially protected groups, such as children.

This investment will be subject to a continuous assessment of the quality of implementation and the achievement of functional objectives, in order to assess at an early stage possible deviations consideration the presence of specially protected groups, such as children from expected future performance or to implement corrective measures.

Box 8. Irish Basic Income for the Arts

The Department of Tourism, Culture, Arts, Gaeltacht, Sport and Media is currently undertaking a pilot [Basic Income for the Arts](#). Applicants were randomly selected into a control and treatment group, where the treatment group are provided with a basic income, in addition to their normal earnings for a three-year period. Both economic and social outcomes are being examined as part of the research. These include hours worked in the arts, hours worked in other sectors, monthly expenditure on arts practice/research, wellbeing indicators, satisfaction with certain aspects of participants lives, deprivation etc. This approach was selected as randomised control trials produce the more robust analysis, and pilot programs are cost effective ways to generate meaningful results when social investment may otherwise be prohibitively expensive, given a lack of evidence

Matching methods

Matching methods select a control group by matching each treated unit with at least one untreated unit that is as similar as possible to the treated one, based on a wide range of characteristics (such as contextual indicators) covered in the data set available (so-called observable characteristics). Therefore, the quality of the control group hinges on the richness of the data available and a large sample size. The most famous of these methods is the **Propensity Score Matching** (PSM). In this method the matching is performed on an index, the Propensity Score, which is defined as the probability of self-selecting into the intervention, for given different individual observed characteristics.

Key requirements for matching methods:

- availability of information on the observable characteristics influencing the selection into the treatment and influencing the outcomes, for both treated and untreated units;
- observable characteristics should be measured before the assignment into the treatment;
- relatively large sample sizes for treated and – in particular – untreated group.

Box 9. Example on propensity score matching: the evaluation of a European Social Fund (ESF) funded intervention in Flanders (Belgium), 2015-2018 (Canzian et al 2020)

The “Work Experience for Young Persons” (WIJ) program is targeted at low-qualified young unemployed with the aim of facilitating their entry to the labour market, through an intensive guiding and counselling activity. Individuals registered at the public employment service were selected and assigned to the treated group using criteria based on age, level of education and previous work experience. However, given financial restrictions not all young individuals registered at employment services were assigned to the intervention, and many of them remained untreated. The analysis aimed at evaluating the impact of WIJ in terms of probability of being employed or re-entering education for young unemployed, some months after the counselling activity was finished. The evaluation of this intervention was done using PSM since the key data requirements were met. First of all, information was available for a good number of observable characteristics, including relevant demographic variables (gender, nationality, date of birth), level of education and previous work history (including months spent in unemployment, employment or inactivity since graduation, number and type of contracts). Second, conditional on these observable characteristics available, it was possible to estimate well individuals’ participation in the program. Finally, a huge number of untreated units were available, sharing similar characteristics with the treated, so that it was possible to select several control units for each treated.

Source: Better Regulation toolbox (2023)

Box 10. Luxembourg private temporary auxiliary contract “Contrat d’auxiliaires temporaires privé” (CAT-PR)

This contract is a paid internship in the private sector, designed to provide young jobseekers (aged 30 or under) practical training to help them integrate into the labour market. The remuneration is generally 100% of the minimum statutory wage “Salaire Social Minimum” for unskilled workers. The state contribution to financing the measure is a reimbursement of 50% of the basic allowance and the full employer’s share of social security contributions.

To measure the returns of this measure, the methodology used is propensity score matching on administrative data supplied by the Public Employment Service (ADEM) and the General Inspectorate for Social Security (IGSS).

The first indicator used to measure the effectiveness is related to the **integration into the labour market**, and answers the following question: Do active labour market policies increase the probability of integration into the job market?

In a second stage, the evaluation was completed by an indicator relating to **the quality of integration** and, more specifically, to the risk of unemployment recurrence: For those integrated, do the measures impact the risk of returning to unemployment?

Below the results:

Chances of integration: 6 months after leaving the CAT-PR, 73% of those who benefited from this measure found a job. Had they not benefited from this measure, only 57% would have entered the job market. **CAT-PR is therefore 16 points more effective.**

Risk of unemployment recurrence: individuals who have benefited from a CAT-PR are less likely to experience a new episode of unemployment in the 18 months following insertion than those who did not benefit from the measure before finding employment (**-13 percentage points**).

Regression Discontinuity Design (RDD)

This approach can be adopted when participation in the intervention is determined by a rule or a threshold on the value of a continuous variable, called the forcing variable or the running variable. Examples of forcing variables are income, age, or a test score. Many spending programs have assignment rules based on thresholds: social benefits based on income, scholarships based on grades, employment programs based on age, etc.

It is common to distinguish between the case when the assignment rule to treated and untreated is applied with no exceptions, called sharp RDD, from the case when exceptions to this strict rule can happen; this latter case is called fuzzy RDD. Fuzzy RDD is technically linked to Instrumental variables (IV), see below.

Key requirements for RDD:

- The probability of participation in the intervention needs to jump at a threshold value of a continuous score (i.e. age, income etc.).
- Enough units (i.e. individuals) should be observed on both side of the threshold (treated and untreated).

Box 11. Example of RDD: the impact of longer unemployment benefits on time in unemployment (Lalive, 2008)

Austria implemented in the late 1980s and early 1990s a program that extended the maximum duration of unemployment benefits from 30 weeks to 209 weeks of unemployed benefit duration to individuals aged 50 years or older who have been living in certain parts of Austria. Given that the program was strictly targeted with respect to age and region, a sharp RDD on age and distance to regional boarder is a natural choice to evaluate this policy. Sharp discontinuities in treatment assignment at age 50 and at the border between eligible regions and control regions identify the effect of extended benefits on unemployment duration.

Source: Better Regulation toolbox (2023)

Differences in Differences (DiD)

The method 'differences in differences' compares the changes in the outcome variable over time between treated and untreated units. Provided that (i) in the pre-intervention period, the outcome variables in the treated and control groups follow similar trends (the 'parallel trends' assumption), and that (ii) the factors explaining differences in outcomes between the two groups are constant over time, differences in the trajectory of the outcome variable between the two groups can be attributed to the participation in the program only. This approach requires repeated observations over time on both groups (so-called longitudinal data or repeated cross sections).

Key requirements for DiD:

- The policy, targeting only a group of units (the treated), should become effective at a certain point in time; this point in time should be known.

- Data for the outcomes of interest should be available both before and after the policy adoption, for both the treated and control group.
- In the pre-treatment period, trends of the outcome variable in the treated group and in the control group should be parallel. This may be checked if a longer time series of the outcome variable before the policy is implemented is available.

Box 12. Luxembourg childcare service voucher “Chèque-Service Accueil” (CSA)

The evaluations of the childcare service voucher (CSA) and the reform of social offices (OS) are part of the mid-term evaluation of the Europe 2020 strategy launched in 2010. At national level, these two schemes have been selected as measures to help achieve two of the five objectives on which the Europe 2020 strategy is focused: employment (increasing the activity rate of women) and poverty reduction. In what follows, the focus is solely on the evaluation of the CSA and its possible impact on the female activity rate.

The introduction of the CSA in 2009: a scheme of partial gratuity and reduced parental financial participation favoring beneficiaries' access to professional educational services; in this case **childcare**.

To measure the results, difference-in-differences method (graphical representation and difference-in-differences regression model) on administrative data is used.

The indicators used to assess the measure's effectiveness include average working hours per month for mothers in couples and for single mothers whose youngest child is aged under 13.

The results show that between 2004-2008 and 2009-2014 (before/after the introduction of the CSA), the average working hours per month of mothers in couples whose youngest child is aged under 13 rose from 85.3 hours to 99.4 hours, an increase of 16.5%, **of which 2.3% is attributable to the introduction of the CSA**. Between 2004-2008 and 2009-2014 (before/after the introduction of the CSA), the average working hours per month of single mothers whose youngest child is aged under 13 rose from 125.1 hours to 127.9 hours, an increase of 2.2%, **of which 1.9% is attributable to the introduction of the CSA**.

Qualitative methods for evaluation

Quantitative methods, like the ones presented above, answer the question “Was there a causal effect of the program or regulation, when did it occur and what is its size?” Once the effects have been estimated, qualitative methods can provide useful information to understand the mechanisms at work behind the estimated impact, also especially when the effects are estimated to be null. This mixture of qualitative and quantitative methods is called the mixed method approach.

Qualitative methods include, for example, **focus groups** or **in-depth interviews** with selected beneficiaries; these can provide the context and help explain the obtained quantitative results, getting insight on the mechanisms underlying the estimated effect of the intervention. They can also provide information on the implementation process, which can be different from how the intervention was initially planned. Any lack of success of the policy intervention could be due either to design failure, when the proposed policy design was not appropriate to achieve the intended objectives in the particular context or implementation failure, when the project was not implemented as originally planned. These considerations are important to guide re-design of future policies.

Mixed methods incorporate qualitative contextual analysis examining the influence of external factors (via the use of participant interviews or focus groups) with process analysis (via similar qualitative techniques applied to the programs' organisational processes) to assess the process of project implementation and how this affected program outcomes and the estimated impact. Some of these methods can offer insights while the intervention is still ongoing, i.e. during the monitoring phase, and they may help to re-target the intervention or to improve some of its

features, even before impacts are estimated using quantitative methods. Qualitative techniques at the beginning of the intervention plan can help to target specific groups of the population, which are not registered in official administrative data and may therefore be not easily reachable through regular surveys' techniques. For example, some marginalised social groups, such as the homeless or irregular migrants cannot be found in registry data, do not appear in official statistics; they can probably be only reached via in-depth field interviews. Similarly, information about some sensitive topics (for example, domestic violence) could be obtained more easily with this kind of interviews, rather than with standard survey data or data from official registries.

Box 13. Qualitative methods as complements ESF evaluation for Flanders (Belgium) (Canzian et al 2020)

This intervention was evaluated using matching techniques, which compared treated individuals to similar non-treated ones. The results show that people participating in this intervention did not show higher employment probabilities than similar people not participating. However, very little was known about the selection process (how case-workers decided which individuals would be assigned to the intervention, once the age and education requirements were met) and on the real contents of the interventions. Qualitative interviews with the case workers can help shed more lights on the selection process. In addition, focus groups or interviews with some of the participants, or qualitative interviews with some of the NGOs in charge of providing the support could provide insights on why the intervention did not bring the expected results. However, qualitative data alone would not have answered the question "Has this intervention worked?"

Source: Better Regulation toolbox (2023)

Ex post macro statistical approaches

Estimates of impacts and returns can be derived from statistical analysis of **country-year data**, such as fixed-effect (FE) panel regression models that are widely applied. They are commonly used in dynamic analysis of macroeconomic returns of social protection expenditure, including **economic growth, economic cycle, and inequality reduction**.

These models are useful for cross-country analyses, as they allow to control for a higher set of observed and unobserved dimensions and heterogeneity by including cross-level interaction effects. While these approaches tend to allow to account for a wide set of aspects at play in a very broad framework, they generally do not allow to identify causality or the channels at play. They use panel data, and – compared to standard pooled Ordinary Least Squares models – allow for reducing the bias from unobserved variables (see Collischon and Eberl A. 2020). To tackle both **unobserved heterogeneity and endogeneity** problems, **System Generalized Method of Moments (system GMM) is generally used** (Arellano & Bover, 1995). Nevertheless, the choice of regression may also be suboptimal (see box and 4 and Clark & Linzer, 2015).

Box 14. macro statistical approaches

Simple regression methods are used to analyse panel data as well but are more widely applied to analyse **cross-sectional data**. Many studies, especially when analysing a single function of social investment or effects on specifically defined outcomes **apply multivariate or bivariate** modelling techniques. Different types of OLS - Ordinary Least Squares or logistic regressions (the latter in case of binary outcome variable) are most widespread. While those can **show interesting correlations**, causality cannot be statistically inferred via those methods. It depends on how comprehensively the study includes relevant explanatory variables and takes necessary precautions to eliminate the endogeneity risk. So-called **instrumental variables**³⁵ can be included to minimize it. Moreover, as the variables on socio-economic data and social protection expenditure are often quite highly correlated with one another, this further

³⁵ An instrumental is a variable that helps identify the causal relation between a dependent variable "Y" and an explanatory variable "X" if it correlates with X and only correlates with Y through X. For example, if "Y" is legal consumption and "X" is illegal consumption, then permissiveness indicators may serve as instrumental variables that correlate with X and only with Y through X.

complicates identifying statistical links. Due to these difficulties, it is not surprising that most studies do not control for potential reversed causality between the variables of interest.

Multi-level models are used to conduct the analyses on **more than one level of data**, to compare the analysis both between and within groups of variables. Such models can be developed for both longitudinal and cross-sectional data. For this reason, they are quite often applied in studies that compare different **social outcomes, e.g., poverty, health, or micro-level employment**, that are reported in **cross-sectional datasets**. This is the case, for example, for part of the EU-SILC data that are only provided as cross-sectional. Multi-level models are also known under different names such as random effects models. Multi-level models require relatively high numbers of observations for each group for reliability (for example, at least 20 observations per country) because with few observations the estimation of the random effect depends more on the abundant number of observations from other groups. Another way to trace cross-country differences can be clustering applied within other standard modelling techniques.

... including frontier efficiency analysis

A few studies provide insights about possible methodologies that attempt to capture social investment effectiveness and efficiency through so called frontier efficiency analysis, which can provide very aggregated results (though without allowing to identify causalities) and are very data intensive (which can also be very sensitive to the actual data that is embedded in the analysis). A discussion of these methods is available in the Social Protection Committee 2015 report. While Frontier analysis has been used relatively widely in the context of education (see for instance Canton and al 2018 or Thum-Thysen and al 2021) it has less frequently been used in the context of social investment since for instance Coelli and al (2008).

Valls Fonayet et al. (2020) analyse the efficiency of social expenditures in poverty risk reduction in the EU28 over the period 2007-2015. The estimates relate to the level of poverty reduction (AROP rate) achieved in link with the social expenditure as a share of GDP. Two indicators are developed: an average frontier to estimate the expected poverty reduction rate of a given country, and the efficiency frontier, an ideal value for the same combination and level of social expenditure. The findings indicate that while higher levels of social protection spending are associated with lower poverty rates, it is not necessarily the biggest spenders that are most efficient. The initial inequality and poverty levels affect the results and diminishing marginal returns can be observed: with high initial poverty rates, social spending brings more prominent results, while reducing poverty further proves more costly. Regarding the types of social expenditure the analysis shows high positive correlation between family/children, sickness/healthcare, and housing spending and efficiency measures.

Buljan et al. (2019) look at the efficiency of public expenditure with regard to health and education (as well as public administration). They rely on Data Envelopment Analysis (DEA) to calculate the efficiency frontiers (and panel Tobit regressions to establish efficiency determinants). The analysis is based on data from 2002-2014 of 15 EEESE countries (11 Eastern EU Member States plus Serbia, Macedonia FYR, Bosnia and Herzegovina, and Montenegro). The selection of the group of countries is based on requirements of DEA programming, which is highly data-sensitive and requires the sample to be highly homogenous (which is why Western EU countries are not included in the analysis). Their findings indicate that level of economic development, education, economic activity, as well as quality of institutions and demographics affect the efficiency of public expenditures.

... and micro-macro approaches

Social Accounting Matrix (SAM) is a framework which can be applied at both micro and macroeconomic levels. It is widely used in economic modelling, as it provides systematic data regarding different economic system inputs and outputs. Such accounts cover the entire economy

and are usually fed into computable general equilibrium (CGE) models. While the macroeconomic simulations, because of the high level of generality, often do not capture returns of social protection expenditure, CGE can be also extended to include microsimulations and address some shortcomings of the latter. Such micro-macro modelling is a promising approach of estimating social protection returns.

Box 15. Social Accounting Matrix (SAM) and computable General Equilibrium Models (CGE)

SAM (social accounting matrix) brings together macroeconomic data (such as national accounts) and microeconomic data (such as household surveys) within a consistent framework and is used for monitoring and explaining the major developments in government expenditure and their links to macroeconomic developments, as well as a basis for macroeconomic modelling (e.g. to calibrate computable general equilibrium (CGE) model parameters).

SAM and CGE have been used for estimating the effects of child benefits (government's 500+ program) on consumption and investments in Poland, for estimating the distributional impact of the 2009 financial crisis on households in Latvia and of several social policy reforms undertaken in response to the crisis. Outside of EU, it was applied for capturing the impact of the social cash transfer program on school attendance and income distribution in Mexico or the impact of the largest cash transfer systems on tackling poverty in Africa.

SAM is an efficient and simple way to record economic transactions. SAMs usually serve two basic purposes: a) as a comprehensive and consistent data system for descriptive analysis of the structure of the economy and b) as a basis for macroeconomic modelling. The construction of SAMs has been described e.g. by Alvarez-Martinez and Lopez-Cobo (2018) based on WIOD (World Input-Output Database) for EU-27 countries, or by Gesualdo and Rosignoli (2009) at the regional level for Italy. Moreover, SAMs could be also constructed at the local/household level to further estimate the impacts of cash transfers on local economies (Taylor, 2012).

SAM was introduced for the first time in the national accounting practice of a developed country in the **Netherlands** in 1994 (Bos, 2008). The SAMs – used by the Netherlands - show the breakdown of government expenditure by function into price and volume changes. For the various (underlying) transactions, prices and volumes can be readily defined. For example, the volume for compensation of employees is the number of full-time equivalents and the price change can be derived as the residual change in the compensation of employees. The volume for social benefits is the number of social benefits. The SAM standard table – used by the Netherlands - starts with the level of government expenditure as a percentage of GDP and then decomposes its change over time into volume and price changes in comparison with those of GDP. Such an overview is very useful for **monitoring, and explaining the major developments in government expenditure and their link to macroeconomic developments** (e.g., wage rate increases, changes in interest rate and changes in the number of social benefits due to general changes in the labour market).

Up-to-date SAM is usually used to calibrate computable general equilibrium (CGE) model parameters³⁶. The **CGE model** reproduces the structure of the economy in its entirety including productive sectors, goods, households, factor inputs, trade and the government, where the nature of all existing economic transactions among diverse economic agents is quantified. The analysis of the effects of social expenditure on poverty/inequality or inclusive growth can rely on tools than integrate microsimulation techniques within CGE. While CGE models focus on the sectoral, macro and price effects of major policy reforms, they generally fail to adequately capture distributional impacts (Cokburn et al., 2014). On the other side, microsimulation techniques focus on the household- and individual-specific distributional effects, but are generally confined to micro reforms as they are unable to model general equilibrium effects – notably on the prices of factors and products, as well as other macro variables – of macro reforms (Cokburn et al., 2015).

Standard CGE models have been used in a number of studies estimating the effects of social protection expenditure at the country level in the EU Member States. For example, the analysis by Górska (2020) [56] used SAM and CGE for **estimating the effects of child benefits (government's 500+ program) on consumption and investments in Poland**. The analysis showed that the government spending resulted in increased revenues and that the program resulted in a positive increase in consumption by 16.64% and in investments by 19.03%.

At the regional level, Viccaro et al. (2020) [57] showed that a two-region SAM, combined with microeconomic information from EU-SILC can be used to show that **proper design of the policy could lead to a decrease in poverty**. The aim of this study was to assess the impact on the income distribution of a social card funded by oil royalties to support the income of poor families in the Basilicata region, **Italy**. The simulation, based on micro-data, enabled an

³⁶ As a SAM is formed to calibrate the model parameters, the structure of a SAM is determined by the structure of the CGE model regarding the number of agents and commodities represented in the model; De Bruin, K., Yakut, A., M (2019). Construction of the Energy Social Accounting Matrix for Ireland, ESRI Survey and Statistical Report Series no. 78.

assessment of the direct impacts of alternative income support systems on the income levels of potential beneficiary households. The analysis was then extended to include also the indirect impacts of income support, taking into account the structure of the region's economic system. This result was obtained by associating the microeconomic dataset to a model of the regional economy based on a SAM. The availability of micro data improved the assessment of impacts on inequality and poverty as opposed to the simple use of the multi-sector model (as in standard SAM analysis). On the other hand, the use of the SAM-based model made it possible to quantify even the indirect impacts on household incomes generated by income support inside Basilicata and in the form of inter-regional feedbacks. This allowed the authors to consider also distributive asymmetries present in the economy (as in Basilicata) that could not be captured by a simple simulation on microdata. The limitation of the approach lies in the fact that the matrix of SAM multipliers that were used in the simulations, reflects the standard assumption of this family of models that are not designed to account for income and substitution effects in consumption.

The top-down CGE approach has been also used for **estimating the distributional impacts of the 2009 financial crisis on households in Latvia and of several social policy reforms undertaken in response to the crisis** such as raising the eligibility threshold for guaranteed minimum income provision (meaning that more people could benefit from it), withdrawing the existing Family State Benefits program and replacing it with targeted cash transfers to the poor, increasing household incomes of unemployed people through augmented unemployment benefits of public works provision (Ajwad et al., 2012) [58]. To measure the distributional impacts of the financial crisis, aggregate shocks to employment and GDP were linked to individual households by combining aggregate information with household survey data. The objective of microsimulations was to estimate Latvian income distribution under different scenarios. Furthermore, the impact on poverty and inequality has been simulated. Both formal and informal employment was considered in the study. The simulations indicated that these measures may have helped cushion the crisis' impacts for some of the hardest-hit households. However, the scale of the crisis exceeded the ability of programs, launched under the Emergency Social Safety Net, to offset its negative impacts. As a result, poor and near-poor people continued to suffer considerable hardship.

Distributional impact assessments (DIA)

An alternative approach to CBA is Distributional impact assessment (DIA)³⁷. This entails a quantitative analysis providing **an assessment of the distributional effects of policies and reforms, hence of income impacts** of reforms and investments **on the income of different socio-economic groups**, including in terms of income inequality and poverty. The analysis focuses on overall income effects rather than on broader equality of opportunities (e.g. in education, or access to the labour market), as these returns are easier to estimate. It thus can help to **elaborate estimates of returns on social investments**.

DIA typically provide estimates of the impact of taxes and social protection and social inclusion reforms and of publicly provided in-kind services. This for instance was useful during Covid-19 or the energy price crisis, including to assess redistributive impacts of in-kind benefits in healthcare or education. Integrating in-kind benefits makes the analysis more comprehensive, while contributing to highlight the potential social and economic returns of this type of social investments. DIAs are of particular importance now when the EU addresses many long-term challenges (demographics, green, digital) and short-term shocks and governments are putting in place related reforms. Recent Eurobarometer survey points at inequalities as one of Europeans citizens' biggest concern (81% of the EU population). Systematic use of DIAs allows to better design measures together with the planned reforms. DIAs can help achieve the EU poverty reduction target of lifting at least 15 million people out of poverty by 2030.

The box below provides an illustrative list of examples of studies using distributional impact assessment.

³⁷ See [Distributional Impact Assessment - Employment, Social Affairs & Inclusion - European Commission \(europa.eu\)](#) in particular the [Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions on Better assessing the distributional impact of Member States' policies](#).

Box 16. Examples of DIA studies

Multiple studies illustrate the various impact of policies across EU Member States. Avram et al. (2012) use EUROMOD to compare distributional effects of fiscal consolidation measures introduced in nine EU countries in the wake of the financial crisis of the late 2000s, finding that the burden of policy changes is shared differently across the income distribution of the countries' populations. The elements of reform - increases in personal taxes, limiting expenditure on cash benefits and reducing public sector pay – led to persons with higher income losing a larger proportion of income than the poor in Greece, Spain, Italy, Latvia, Romania and the UK, while in Estonia the poor lost a higher proportion of income than the rich. The study also assessed the effects of all the tax-benefit policy changes introduced in the period 2008-12 and found that, while generally reducing incomes, the scale of the effects was smaller and the distributional impact was similar or less progressive than the austerity changes looked at in isolation. The evidence highlights that the broader national tax regulation context and the specific income distribution of a population are moderating factors to policy change effects. The distributive effects of the financial crisis and subsequent austerity measures were also discussed based on EUROMOD simulations by Matsaganis and Leventi (n.d.). Barrios et al. used EUROMOD to analyse the fiscal and distributional impacts of tax expenditures related to public and private contributory pension schemes on both contributions and pension benefits in all EU Member States (Barrios et al., 2020). The study found that pension-related tax expenditures have a sizeable impact on tax revenue. The authors argue that in the context of economic downturn and aging of populations, more needs to be done to improve the efficiency and long-term sustainability of pension schemes in Member States.

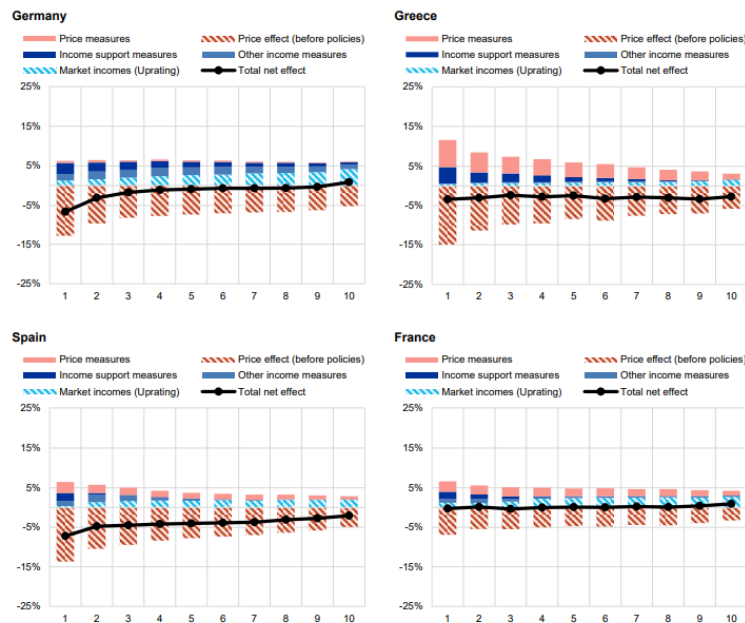
To estimate the labour supply effect of cash benefits, the first step is to quantify monetary incentives. The most often used approach is to analyse the difference in incomes when out of work and when taking a job. Some analysed the impact of a newly implemented policy, e.g. Christl and De Poli (2021) the "Neue Sozialhilfe" (New Social Assistance) in Austria. Colombino et al. (2008) focused on the effect of 10 hypothetical tax-transfer reforms that were alternative versions of a basic income policy in European countries. O'Donoghue (2011) simulated the participation replacement rate and the participation tax rate highlighting the important role of household composition and incomes in Europe. Xavier, Gasior, and Makovec (2017) provide a description of the distribution of short- and long-term participation tax rates and marginal effective tax rates. Immervoll et al. (2007) compares the effects of increasing traditional welfare to introducing in-work benefit by calculating marginal and participation tax rates. Kalíšková (2015) compared the effect of tax-benefit policies on female labour supply in 26 European countries. Bargain, Orsini, and Peichl (2011) beyond the labour supply response attempted to compare elasticities in the selected countries.

Source –Ecorys (2024).

A recent ECB paper analyses the distributional impact of high consumer inflation in the euro area and government measures to compensate households in 2022. The study uses the tax-benefit microsimulation model for the European Union (EUROMOD) with microdata as the input – EU statistics on income and living conditions (EU-SILC) and household budget surveys (HBS) – to quantify the distributional impact of inflation, income support measures and measures aimed at containing prices. The analysis confirms that purchasing power and welfare were more severely affected by the 2022 inflation surge in lower-income households than in higher-income households. Fiscal measures compensated households for about a third of their welfare loss, though with significant differences between countries. At the same time, fiscal measures closed around 60% of the inequality gap between lower and higher income households. Most fiscal measures were not particularly well targeted at low income households, resulting in a higher than necessary fiscal burden to cushion the distributional impact of the inflationary shock. The figure below presents the results for some Member States (Amores et. al., 2023).

Price and income effects based on households' welfare in the euro area countries

(percentage change in equivalised disposable household income, per decile)

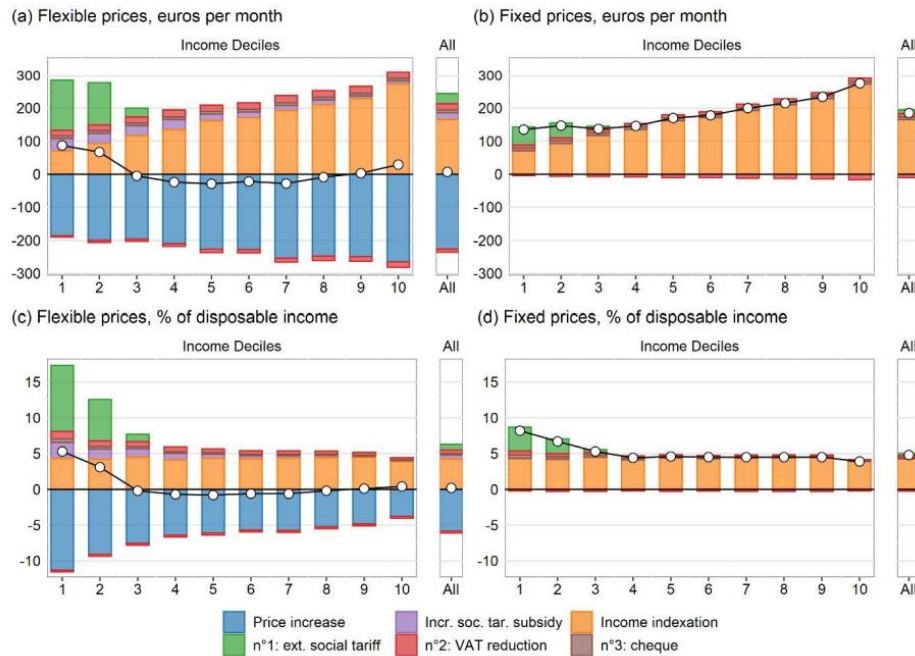


Source: 'Inflation, fiscal policy and inequality. The distributional impact of fiscal measures to compensate consumer inflation'. ECB Occasional Paper.

A similar study by Capeau et. Al (2022) analyses the distributional impact of the price shock and the government measures together with the automatic indexation mechanism for Belgium. The main findings are:

- Households with new or variable price contracts in the middle-income deciles are the ones least compensated by the compensating measures. That said, they too – on average – see a large part of the shock compensated.
- The measure that compensates the most for the loss in purchasing power, with a wide margin, is the automatic indexation of incomes. It compensates the higher incomes more than the lower incomes.
- At least in the short term, there is an important difference in how households with fixed and variable price contracts are affected. Households with fixed price contracts do not experience any negative impact on their purchasing power as long as their contract has not expired, while benefiting from automatic indexation, and from two of the three new government measures: the extension of the social tariff to a broader group and the heating cheque.
- The VAT reduction is on net slightly beneficial for low incomes, while slightly negatively impacting higher incomes, but overall very small.
- The social tariff and the broadening of the eligibility criteria for this tariff are key in compensating the bottom deciles.
- The VAT reduction's effect on purchasing power seems rather small, especially when considering the secondary effect on the indexation of incomes.

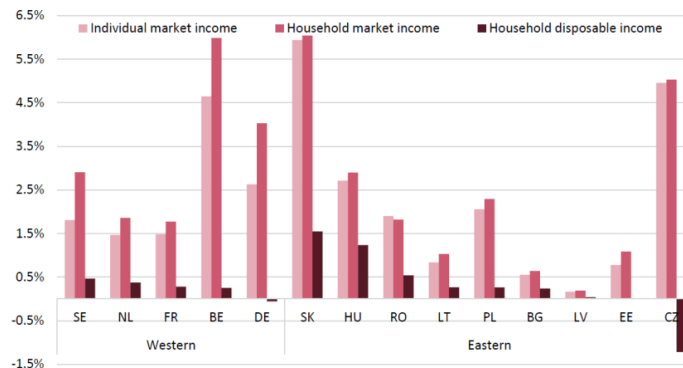
Figure 4: Change in disposable income and in expenditures per measure (euros per month and % of income)



Source: Capeau, B., Decoster, A., Güner, D., Sheikh Hassan, N., Vanderkelen, J., Vanheukelom, T. en Van Houtven, S. (2022), 'We are all facing the same storm, but not all are in the same boat', Discussion Note, KU Leuven Department of Economics.

Another recent study analyses the effects of robot penetration on household income inequality in 14 European countries between 2006–2018, a period marked by the rapid adoption of industrial robots. Automation reduced relative hourly wages and employment of more exposed demographic groups, similarly to the results for the United States. Using robot-driven wage and employment shocks as input to the EUROMOD microsimulation model, we find that automation had minor effects on income inequality. Household labour income diversification and tax and welfare policies largely absorbed labour market shocks caused by automation. The figure below shows that transfers played a key role in cushioning the transmission of these shocks to household incomes (Doorley et.al., 2023). The study illustrates the potential of DIA to pro-actively address the digital transition.

Figure 5. The effect of automation on income inequality using various income concepts



Notes: The figure shows the change in Gini Index due to automation where income is defined as (i) market income at the individual level (ii) equivalised market income at the household level and (iii) equivalised disposable income at the household level. In Eastern and Western Europe, countries are ordered in decreasing order of the total change in the Gini Index due to automation. Data: EUROMOD, EU-SILC.

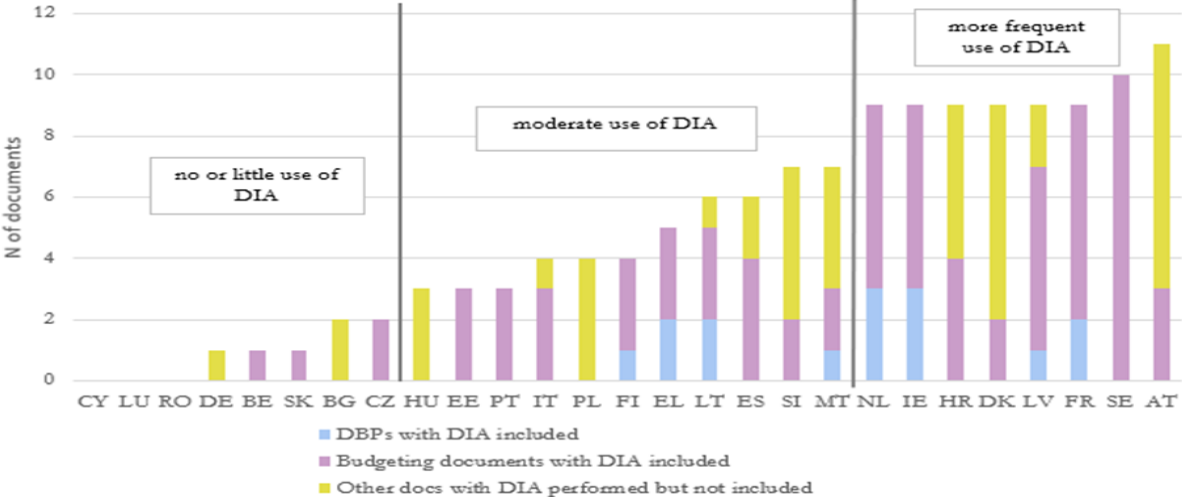
Source: Doorley, Karina & Gromadzki, Jan & Lewandowski, Piotr & Tuda, Dora & Van Kerm, Philippe, 2023. [Automation and Income Inequality in Europe](#), IZA Discussion Papers 16499, Institute of Labor Economics.

At government level, DIA practices are generally in place in most EU Member States typically in both economic and finance ministries and social ministries (though not necessarily disseminated to the wider public) and used among others in Drafting Budgetary Plans.³⁸ They are often used to estimate impacts and returns in case of specific reforms. They can also present the effect of a package of policy changes, as some interactions may be missed if analysed separately. By allowing the comparisons of different policy options and testing alternative policy design combining several features, DIA helps identify the need to target policy measures and how to maximise returns, also in terms of income distribution. DIA can be used both ex-ante and ex-post: ex-ante it can contribute to the design of measures, while ex-post it can feed into policy evaluation and eventually help identify further measures. DIA techniques generally rely on microsimulation models that account for the tax-benefit systems in place and the related policy changes. They can also account for behavioural modules or macroeconomic developments (e.g. demographic trends or macroeconomic feedback loops) that enrich the model and improve their accuracy. An important advantage of DIA, based on micro-simulation models, is that they usually allow for a simultaneous assessment of social and fiscal impacts of social investments. As such they can also contribute to a cost-benefit analysis.

In terms of use of DIA across member states, significant differences emerge across member states, as illustrated in the figure below:

- Some MSs use systematically DIAs for their budgetary (and reform) measures, some rather sporadically or ad-hoc. Only very few MSs do not appear to conduct any DIA analysis.
- Some do DIAs and publish full results in their national budgetary documents.
- Others carry out the analysis without making it public.
- Others present generally only a descriptive analysis of the distributional impact of their policies.

Figure 6. DIA practices in Member States



Source: Study on Distributional Impact Assessment by the University of Milan, [Study on distributional impact assessment - Publications Office of the EU \(europa.eu\)](#)

³⁸ See [for instance the Commission Staff Working Document Accompanying the Communication.](#)

Based on the questionnaire on DIA progress in 2023 by the European Commission, 12 Member States reported with overall some developments:

- 5 MS report developments in DIA methods (BE, FI, IE, IT, NL) ;
- 2 MS report developments in integration of DIA in policy making (IE, LV) ;
- Other MS reported: Plans for developments (CZ) or organisational changes (DK).

In the following box, we describe examples of DIA uses in several different Member States of the IWGSI. These examples sometimes share a commonality in developing and utilizing models based on the European model EUROMOD. Additionally, the issue of the data used remains a challenge.

Box 17. National examples of DIAs

Some EU Member States have integrated DIAs into their social policy planning.

In **Austria**, the Federal Organic Budgeting act (Federal Organic Budget Act 2013, Sec 17, English translation) stipulates that the member of the Federal Government or the budget holding body (e.g. the President of the Court of Audit), in whose area of impact the draft of a legislation was prepared or the project was planned, must attach an impact assessment. (Motions for draft legislation submitted by members of Parliament are not subject to mandatory Outcome Oriented Impact Assessments). The responsibility for carrying out outcome-oriented impact assessments is therefore aligned with the federal body which is responsible for the project concerned. The Austrian “Outcome Oriented Impact Assessment” („Wirkungsorientierte Folgen-abschätzung“), which requires to assess legislative proposals and other projects (e.g. large social funding programs) with regards to its impacts on several different target groups, such as impacts on the labour market and working conditions, on people with disabilities or in need of long-term care and including the AROPE-target group (others include: impacts on the labour market and working conditions, on people with disabilities or in need of long-term care and others). When a project is planned an OOIA has to be drafted and submitted to the Federal Performance Management Office. DIA analyses are discussed in parliament and by experts on an ad hoc basis. Evaluation reports are submitted annually to the Budget Committee of the National Council. Impacts on the AROPE-target group can be assessed with the Austrian model SORESI. The Austrian model SORESI builds on EUROMOD, but is adapted for Austria and accessible to the public free of charge and without any further requirements (software, data, etc.). It is accessible to the public via this website: <https://soresi.sozialministerium.at/>. It is based on EU-SILC data and updated annually, as the new data becomes available. For the assessment of complex fiscal and economic effects of tax and transfer reforms, the Fiscal Council also suggests the use of appropriate proven models (microsimulation models, macro models as well as combined models) for the “Outcome Oriented Impact Assessment”.

In **Finland**, the Ministry of Finance, along with relevant ministries, is responsible for assessing the distributional impact on the government budget. They prioritize ex ante overall impact assessments and are continually developing these assessments, often through horizontal projects that go beyond social or employment policies. Finland has extensive administrative data and utilizes microsimulation models for evaluations. However, resource constraints pose challenges, requiring constant prioritization.

Ireland has a long standing and well-established tradition of conducting distributional impact assessments (DIAs), a practice upheld for over a decade. The results of these assessments on potential budget measures are published in advance of the budget. Furthermore, they provide an evaluation of the actual budget measures as part of the Budget day documentation, followed by a more detailed assessment post-budget. The assessment looks at how taxes and welfare impact different income groups, household types, and age groups. It relies on model based on EUROMOD, with ongoing development supported by government funding through the Economic and Social Research Institute. Recent enhancements include the ability to evaluate the impact of non-financial measures such as childcare, medical cards, and free school books. Despite these advancements, challenges persist. One notable challenge is the limited time horizon of the DIA conducted, which only consider current impacts due to constraints in the existing model. Additionally, as the model relies on survey data, measuring impacts on specific subgroups of interest poses difficulty. However, recent improvements involve the incorporation of an indicator focusing on families impacted by disability. Ongoing efforts include collaboration with the National Statistics Office to explore linking census survey administrative data, aiming to extend the model's capability to assess impacts on certain groups

beyond its current scope.

In **Italy**, the practice regarding impact evaluation of reforms, especially in the budget law, is primarily carried out by the Ministry of Finance, utilizing traditional models such as the QUEST model promoted by the European Commission for macroeconomic impact analysis. The Ministry of Labour, on the other hand, does not conduct DIA. However, there is an annex to the budget law known as the sustainable indicators, incorporating 12 or 18 key indicators with projections based on the budget law's measures. While not a direct DIA, it provides an evaluation of the impact on various aspects, including poverty and early school leaving.

In **Lithuania**, the Ministry of Social Security and Labor, in collaboration with the Ministry of Finance, primarily utilizes the EUROMOD tax-benefit microsimulation model for conducting DIAs. The focus is on ex-ante assessments, estimating the net budgetary cost of changes to taxes and benefits, analyzing patterns of gains and losses resulting from policy changes. It allows for the assessment of the impact of reform packages on poverty and inequality, including evaluations by gender and other quality criteria depending on available data. The EUROMOD model has been adapted in Lithuania to perform more detailed simulations, incorporating EU SILC data and developing a prototype based on merged administrative records. The production of administrative input data for the model is transitioning to a new National State Data Management Information System, a significant step in strategic planning, serving as a central national data point by merging various systems. Before the creation of the new system that acts as a single platform for strategic planning management indicators, in Lithuania there were 95 registers and 275 state IS, managed by 134 institutions and served by 166 different operators. Lithuania has implemented distributional impact assessments using EUROMOD in nine areas over the past few years, examining aspects such as the introduction and increase of single-person benefits, changes to the minimum income protection system, pension indexations, minimum wage and tax allowance increases, and alternatives to tax reform. There is a recognition of the need to strengthen the overall distributional impact assessment, possibly extending it to cover the entire national yearly budget, with plans for a joint distributional impact assessment report submitted to Parliament.

In **Luxembourg**, DIA are primarily done on an ad hoc basis, with help of the SPAFIL (Social Policy Analysis File for Luxembourg) microsimulation model. Based on individual administrative data, SPAFIL models social benefits contributions and income tax, estimating the impact of reforms on revenue, expenditure, income gains or losses, and income inequalities. While DIA use is not systematic in Luxembourg's administration, there is a growing perception of its importance in evidence-based decision-making.

When it comes to experience, the **Netherlands** possesses substantial expertise in conducting Distributional Impact Assessments (DIAs), particularly in the lead-up to the presentation of national budget plans. These assessments primarily focus on calculating the purchasing power of various societal groups. However, a notable challenge arises in the inherent bias towards emphasizing groups that are more prominently visible in the assessment results, a concern that experts attempt to address, facing limitations. Another challenge is the tendency for political discussions to predominantly center around purchasing power, neglecting broader aspects such as poverty and inequality impacts. This dynamic poses a dilemma for experts presenting results to politicians. Additionally, in evaluating policy measures, determining the effectiveness of each measure proves challenging due to the intricate web of interconnected and relevant social policies. For instance, when assessing the adequacy of the minimum wage in the Netherlands, households also receive various compensatory benefits for expenses like rent, children, and healthcare insurance. While these benefits contribute to the households all kind of income, gauging their effectiveness in ensuring a decent standard of living remains a complex task. Despite the complexity, efforts are made to separate and assess these factors on a complementary basis.

In **Slovakia**, Distributional Impact Assessment (DIA) involves ex-post empirical analyses of policy impacts that partition the sample to identify heterogeneous effects, especially in cases with available data. For instance, impacts on workers in distinct regions or sectors are estimated separately. However, certain types of distributional impact analysis, such as deconvolution, are not used due to their dependence on restrictive assumptions. In the ex-ante phase, various policy effects are forecasted for different stakeholder groups, including households and employers.

In **Slovenia**, the Ministry of Labour, Family, and Social Affairs has developed microsimulation models related to pensions, the labor market, and social transfers based on the EUROMOD model. They have collaborated with the OECD to enhance their capacities in this area under the Commission's structural reforms support. These models can assess the complementarity of different measures and policies, particularly in social affairs, aiming to improve the efficiency of their measures. The models utilize administrative data when it is possible, addressing challenges related to access, compatibility, and GDPR issues.

In **Sweden**, the consideration of distributional implications in the budget process has been a longstanding practice. Since 1994, an annual report on income inequality and, since 1988, a report on gender inequality have been published. These reports include ex-ante distributional impact analyses. The Division for Economic Policy and Distribution (DEPD) within the Ministry of Finance plays a crucial role in analyzing the distributional impacts of proposed policies in the budget on income inequality and economic gender equality. This includes examining the effects of taxes, transfers, and publicly funded welfare services. The model is available to all government agencies

free of charge, while the Parliament and other users pay a users' fee. External organizations are able to access the code, excluding the data, or can order analyses from Statistics Sweden for a fee. It can also evaluate how regulatory changes affect marginal effects and replacement rates for households (both for specific social and income groups, or aggregated to the societal level). Finally, it evaluates statistics on publicly funded welfare services. While predominantly static, FASIT includes a labor supply model which is able to estimate the effects of tax and transfer changes on the long-term labour supply and the implied long term effect on income distribution. The module contains detailed rules for taxes and transfers, data on income, and several estimated equations based on individuals' characteristics (education level, household type, etc.) that partly describe individuals' preferences for market work, and partly examine the probability of their transitions from non-work to work when the compensation rate changes. The module is also able to consider labour market heterogeneity – for example, it accounts for the fact that different types of households (e.g., single women, single men, cohabitants) are likely to have different work preferences. However, FASIT has limitations. It is mainly static and cannot simulate real-time behavioral reactions to welfare changes. Analyses are conducted on a yearly basis, potentially missing month-to-month impacts of transfer changes. Some policies, especially those related to collective public goods (such as police and defense) or reforms where service users are uncertain, cannot be simulated. Additionally, the reliance on a household survey for indirect tax calculations presents a limitation due to historically low response rates, making some information outdated.

Distributional impact assessment and budgetary plans

At the EU level, the Commission published in 2022 a Communication on Better assessing the distributional impact of Member States' policies and related Staff Working Document. Further support for Member States is provided at the EU level in particular via the maintenance and further development of EUROMOD, assistance under the Technical Support Instrument, Mutual learning to facilitate exchange of best practices among MSs, monitoring of the use of DIA, and continued dialogue on enhancing the use of DIA in the Committees (EMCO, EPC and SPC).

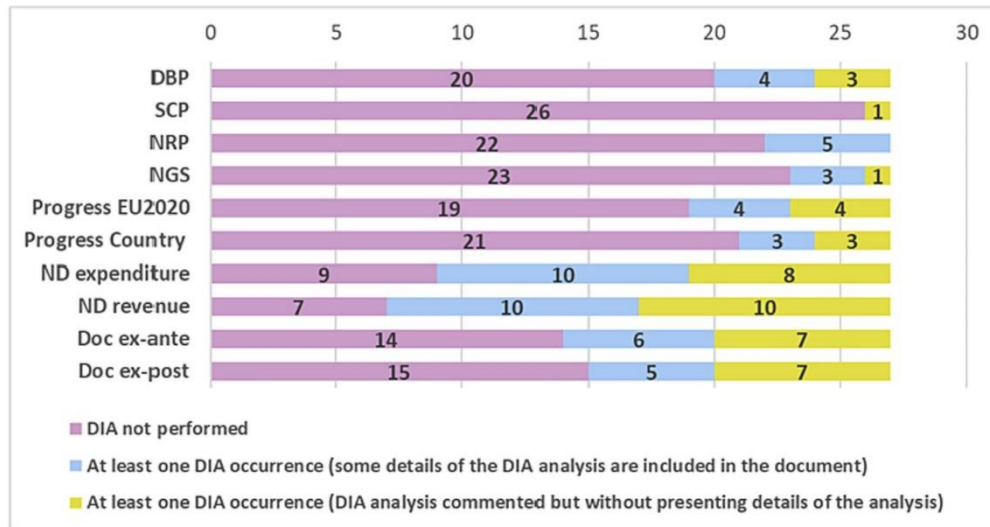
Furthermore, as mentioned above, Article 6(3)(d) of regulation 473/2013 on common provisions for monitoring and assessing draft budgetary plans (DBP) and ensuring the correction of excessive deficit of the Member States in the euro area, stipulates that where possible the DPB should include an indication on the expected distributional impact of the main expenditure and revenue measures.

More recently, the Commission proposal for a Council Directive amending Directive 2011/85/EU on requirements for budgetary frameworks of the Member States, invites the latter to publish information on the distributional impact of budgetary policies and take into account employment, social and distributional aspects in the development of green budgeting (see article 9).

Bazoli et al. (2023) investigated the extent to which EU MSs use DIA in their budget preparation process and what limits the use of DIA, looking at the 2015-2020 main budget documents: Draft Budgetary Plans (DBPs), Stability and Convergence Programs (SCPs) and National Reform Programs (NRPs).

What emerges is that, only seven countries (Estonia, Finland, France, Greece, Ireland, Lithuania and the Netherlands) out of the nineteen Euro Area countries, the EU MSs to which the EU Regulation No 473/2013 applies, feature at least one DIA occurrence in their DBPs. Concerning SCPs, only Croatia exhibits DIA, although results are only commented not presenting details of the analysis; while five countries (Belgium, Croatia, Italy, Malta and Portugal) include DIA in the NRP document. Four countries (Austria, Denmark, Slovenia and Sweden) perform at least one DIA occurrence in the National Growth Strategy document, while eight countries (Austria, Denmark, France, Latvia, Malta, Portugal, Slovenia and Spain) present at least one DIA occurrence in progress towards EU2020 indicators documents, and the same countries, with the exception of Slovenia and Spain, provide DIA analysis in progress in implementing country-specific recommendations documents.

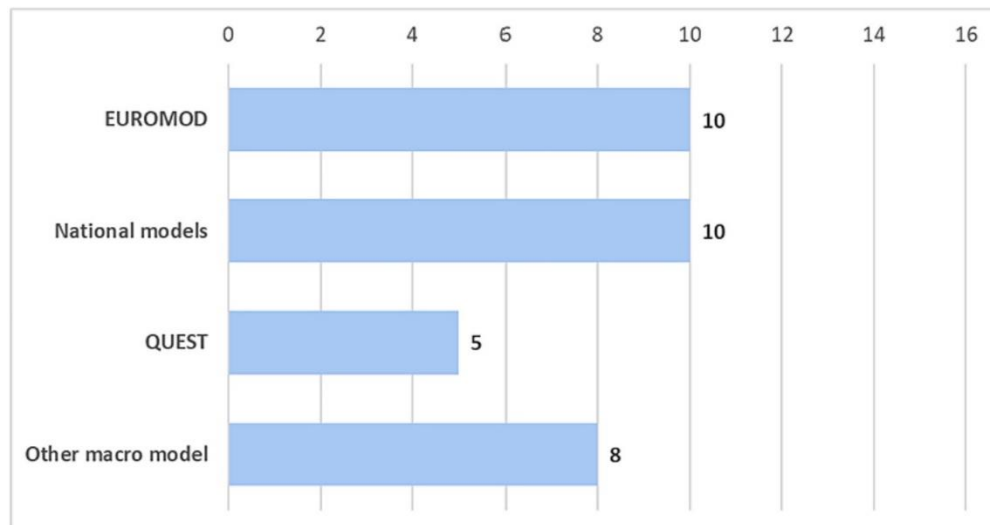
Figure 7. Number of European Union Member States With at Least One DIA Occurrence in 10 Official Budget-Related Documents, 2019–2020.



Source: Bazoli et al. (2023) Equal Focus on Inequality? Approaches to Distributional Impact Assessment in the National Budget Process Across the EU, Journal of Common Market Studies

Turning to the question of how DIA is performed in the MSs, the figure below shows the models used to perform DIA. Two Europe-wide models, EUROMOD and QUEST³⁹, are featured prominently in the responses.

Figure 8. Microsimulation Models and Macroeconomic Models Used to Perform DIA.



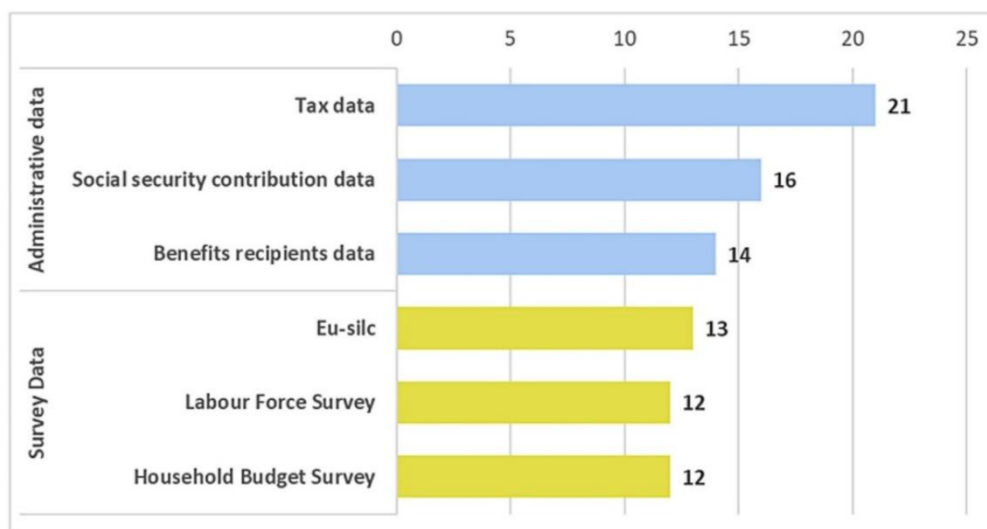
Source: Bazoli et al. (2023) Equal Focus on Inequality? Approaches to Distributional Impact Assessment in the National

³⁹ QUEST serves as the global macroeconomic model utilized by the Directorate General for Economic and Financial Affairs (DG ECFIN) for conducting macroeconomic policy analysis and research. It operates as a structural macro-model following the New-Keynesian tradition, incorporating thorough microeconomic foundations derived from utility and profit optimization. This includes considerations for frictions in goods, labor, and financial markets. See [QUEST macroeconomic model – European Commission \(europa.eu\)](https://ec.europa.eu/economy_finance/quest).

Distributional impact assessment, data used and implementation

Regarding the data utilized for Distributional Impact Analysis (DIA) across the EU, it is common to use both surveys and administrative data. Among these, tax data, followed by data from the EU-SILC, the Labour Force Survey and the social security data, are the most frequently used data sources for DIA. The large use of administrative data during the budget process is not surprising as they tend to exhibit less noise than survey data and tend to be easily accessible to ministry officials. The data for DIA are usually provided upon specific request, and the procedures can differ between survey and administrative data. Survey data are provided by Eurostat or by National Statistical Institutes, while administrative data are requested within the public administration.

Figure 9. Data Used for Producing DIA Analysis.



Source: Bazoli et al. (2023) Equal Focus on Inequality? Approaches to Distributional Impact Assessment in the National Budget Process Across the EU, Journal of Common Market Studies

As stressed in Bazoli et al. (2023) the quality of data significantly affects the quality of DIA. The main problem regards time lags of survey data which negatively affected the capacity of MSs to perform DIA during the pandemic crisis. To overcome this problem, Lithuania, which relies only on survey data, is working in collaboration with the JRC on a project aimed at adjusting the microsimulation model to run on administrative data. The same methodology of complementing Euromod with administration is also being applied, through collaboration with the JRC, in Greece, Romania, Slovakia, Latvia and Spain, while the Federal Social Security of Belgium has produced such a model independently of JRC.

In addition to data quality, several factors pose obstacles to the effective implementation of Distributional Impact Assessment (DIA). The timing of the budgetary drafting process stands as a significant challenge, particularly in many Member States where the Draft Budget Proposal

(DBP) is finalized mere hours before the October 15 deadline. This leaves insufficient time for a thorough and valid DIA, even if the necessary tools, expertise, and data are available. Furthermore, low political awareness, inadequate human resources allocated to DIA, and a lack of comprehensive knowledge about DIA, whether at a basic or advanced level, further hinder its seamless integration into the budgetary process. Establishing a dedicated team with the capacity to perform DIA throughout the budget process is crucial for overcoming these barriers.