



Beyond financial proxies in Cohesion Policy inputs' monitoring: A system dynamics approach

Giovanni Cunico^{a,*}, Eirini Aivazidou^{a,b}, Edoardo Mollona^a

^a Department of Computer Science and Engineering (DISI), School of Science, Alma Mater Studiorum – University of Bologna, Mura Anteo Zamboni 7, Bologna, 40126, Italy

^b School of Mechanical Engineering, Faculty of Engineering, Aristotle University of Thessaloniki, P.O. Box 461, Thessaloniki, 54124, Greece

ARTICLE INFO

Keywords:

Cohesion Policy
Structural funds
Absorption rate
Performance evaluation
Monitoring indicators
System dynamics

ABSTRACT

European Union's Cohesion Policy aims to foster development and reduce disparities among regions by redistributing more than one-third of the European budget. Given the policy's importance and complexity, an elaborated monitoring and evaluation system has been established. While attention has been dedicated to evaluating policy impact, the monitoring of inputs (i.e., allocated financial resources) has been limited to the control of financial dimensions (i.e., funds' absorption rate). As the implementation process entails a sequence of steps, this research explores whether financial proxies alone are adequate to monitor the policy inputs. To test this hypothesis, a system dynamics model is built. Simulations highlight that the absorption rate captures shocks that might occur during the inputs' expenditure with significant delay. To that end, we elaborate three novel operative monitoring indicators (i.e., funds' demand, funds' offer, procedural efficiency), which may overcome the financial indicators' mono-dimensionality and time lags' limitations.

1. Introduction

The Cohesion Policy (CP) is one of the most important policies in the European Union (EU), accounting for the second-largest budget expenditure (European Commission, 2014). CP aims to support sustainable and harmonic progress, especially by fostering economic growth in less developed European regions and sustaining competitiveness of the most developed ones through the redistribution of economic resources (i.e., structural funds). The policy has been active since 1988 and is divided into cycles of seven-year programming periods (Brunazzo, 2016). It is defined and supervised at a European level, while the implementation is performed by the Member States and the local managing authorities (LMAs), such as regions. This distribution of jurisdiction among different authorities is called multi-level governance (Hooghe, 1996; Piattoni & Polverari, 2016). Thus, CP encompasses high complexity (Bachtler & Ferry, 2015; Rogers, 2008; Stephenson, 2013), constituting an open, multi-level, multi-focus, and multi-action system (Lion & Martini, 2006) in which actors' decisions affect the system with considerable delays. However, the actual effects of CP are not homogenous (Bachtler, Fratesi, & Perucca, 2020; Fratesi & Wislade, 2017) and several European regions face difficulties in employing the allocated structural funds (European Court of Auditors, 2018; European Parliament, 2011). This

unreliability of absorption potentially hinders the accomplishment of CP's objectives. Against this background, an articulated and extensive mechanism of monitoring and evaluation system has been developed (European Commission, 2018a; European Parliament & the Council of European Union, 2013), rendering CP "the largest and most evaluated policy in the world" (Naldini, 2018, p. 496). This is unsurprising since evaluation literature has repeatedly showcased and discussed the intricacy of evaluating complex systems (e.g., Hassmiller Lich et al., 2017; Moore et al., 2019; Rogers, 2008).

The CP evaluation and monitoring system has evolved over time due to the contribution of practitioners' direct experiences (Gaffey, 2013; Lion & Martini, 2006) and researchers' efforts (Barca, 2009; Boselli & Tagle, 2003). Therefore, the academic debate on meta-evaluation and monitoring could be crucial for shaping the CP evolution. This debate has a long and vivid tradition; starting from the early stages of CP (Bachtler & Michie, 1995; Mceldouney, 1991), it has proficiently continued over policy cycles (Bachtler & Wren, 2006) and is continuously being stimulated (Naldini, 2018). The discussion has revolved around several aspects of the CP monitoring and evaluation system, indicatively: (i) the role of evaluation units and monitoring committees (Cartwright & Batory, 2012; Olejniczak, Raimondo, & Kupiec, 2016), (ii) the indicators' misuse (Nigohosyan & Vutsova, 2018) and

* Corresponding author.

E-mail address: giovanni@cunico.org (G. Cunico).

<https://doi.org/10.1016/j.evalprogplan.2021.101964>

Received 28 July 2020; Received in revised form 9 December 2020; Accepted 11 May 2021

Available online 31 May 2021

0149-7189/© 2021 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

development (Barca & McCann, 2011; Barca, 2009; Masana & Fernández, 2019; Marchante & Ortega, 2010), (iii) practical lessons in the field (Gaffey, 2013; Lion & Martini, 2006; Perrin, 2011), (iv) the systems' fundamentals (Batterbury, 2006; Perrin, 2011; Saunders, 2011), (v) how it has evolved over time (Gaffey, 2013; Naldini, 2018), and (vi) its formative impact on organisations (Wojtowicz & Kupiec, 2018).

To the best of our knowledge, most research efforts have been dedicated to the evaluation of the CP outputs, outcomes (or results), and impacts, while little attention has been given to the monitoring of the policy inputs, namely the funding resources allocated (i.e., structural funds). In this respect, Masana & Fernández, (2019) identify a lack of maturity and development of local CP monitoring, which is primarily based on a conventional accounting/financial approach. In fact, based on the current regulation, the inputs' monitoring is conducted with financial means and proxies, particularly through the absorption rate of the funds' expenditure. Further actions are generally left to the LMAs and Member States (European Commission, 2018b) and, when they occur, the responsible parts avoid to share them in a systematic and transparent manner. Policy-makers, researchers, and the media often use the absorption rate to assess and monitor the state and success of a CP programme (Aivazidou et al., 2020; Gaffey, 2013). However, the structural funds' expenditure and implementation constitute a long 'pipeline' process entailing a series of steps (e.g., funds allocation, calls' issuing, applications' appraisal, projects' implementation, applicant refunding) that require a considerable amount of time (Cunico et al., 2020; Stephenson, 2016). This research gap appears particularly concerning as the absorption rate measures the inputs' utilisation at the end of the aforementioned pipeline. In addition, this indicator alone might not be adequate for monitoring CP inputs since its mono-dimensionality (Masana & Fernández, 2019), along with the delays involved in the implementation process, might hinder timely and comprehensive understanding of the implementation programme's efficiency. Better comprehension is compelling as CP funds' absorption issues still occur in several European regions (European Court of Auditors, 2018). Thus, a more effective inputs' monitoring could sufficiently predict these shortcomings and support CP agents in taking adequate corrective actions on time. Moreover, the use of exclusively financial dimensions for monitoring can conceal the use of debatable policies, such as retrospective projects' use¹ or regional co-finance reduction² (Corte dei Conti, 2017; European Court of Auditors, 2018), utilised by some LMAs to increment the absorption rate values without increasing the actual amount of local investments. This specific necessity for a 'better' monitoring of the inputs has been already identified by Bachtler & Ferry, (2015, p. 1259): *"the progress of development programmes cannot be meaningfully assessed because of the lack of meaningful and timely information on programme performance"*.

In this respect, this paper poses the following research questions (RQs):

- RQ#1: *Can financially oriented indicators (i.e., absorption rate) alone provide adequate monitoring of the CP implementation system?*

¹ *"Retrospective projects are those which have incurred expenditure from national sources or are completed before EU co-financing has been formally applied for or awarded, i.e. they are financed retrospectively. In the 2014–2020 program period, projects or operations that are physically completed or fully implemented before the beneficiary submits the application for funding are not eligible for EU funding"* (European Court of Auditors, 2018, p. 5).

² CP funds are composed of the European share and the national/regional one. The national and/or regional co-finance reduction refers to the decrease of national and/or regional contribution, respecting the regulatory limits (exceptions can be made in extreme cases) (European Commission, 2013). This strategy is expected to speed up the absorption of EU resources at the expense of the total amount of resources invested locally.

- RQ#2: *Can alternative monitoring indicators go beyond the economic mono-dimensionality of absorption rate and detect problems earlier in the implementation process?*

To respond to the RQs, this paper aims to assess formally the adequacy of the funds' absorption rate as a monitoring indicator by evaluating its systemic behaviour over time after a variety of exogenous disturbances occurred (RQ#1). Then, in case the absorption rate's monitoring effectiveness is insufficient, we explore whether additional operational indicators, developed for the first time, could complement the use of the financial ones to improve monitoring performance (RQ#2). To test these hypotheses, we built a simulation model replicating the implementation process of CP funds of an LMA using the system dynamics (SD) method (Forrester, 1961; Sterman, 2000). Systemic methodologies have already crossed their paths with the evaluation domain, fruitfully integrating enhanced evaluation practices and theories (Gates, 2016, 2017; Hassmiller Lich et al., 2017). This approach allows replicating the operational steps (Olaya, 2015) that occur when an LMA implements the funds, simulates different scenarios, and tests the behaviour of absorption rate in these occasions. In addition, the model structure enables the identification of the key points of the implementation pipeline, which could be relevant to monitoring in case of developing novel operational indicators. Besides, this paper further anticipates to contribute to the current debate on CP monitoring and evaluation, as framed by Naldini (2018), by expanding the focus to the inputs' monitoring for management purposes. In addition, we attempt to showcase that simulation models (e.g., SD models) can effectively support the development and testing of a new-generation set of operational indicators for CP, improving inputs' monitoring and, overall, absorption comprehension.

The remainder of the paper is organised as follows. Firstly, the theoretical background is described to contextualize this research effort (Section 2). The SD approach, the data collection, and the model structure and validation are outlined in Section 3. Then, the simulation results of the absorption rate are presented, while the new original indicators and the related dynamic behaviour are investigated (Section 4). The major outcomes, implications, and challenges are analysed and discussed in Section 5. Finally, we conclude with the key insights, proposing future directions (Section 6).

2. Theoretical contextualisation

2.1. Towards a monitoring and evaluation integrated framework

Although monitoring and evaluation are both management tools which are usually mentioned interchangeably when discussing CP assessment (Perrin, 2011), they constitute two different concepts (Bachtler & Wren, 2006; Gaffey, 2013; Perrin, 2011). Monitoring constitutes the systematic and continuous collection of data for tracking the progress of a specific process while the project activities are on-going. It aims to provide information about the current state of the system and, thus, support the adoption of prompt and informed remedial actions by decision-makers, if needed. On the contrary, evaluation is the periodic appraisal of project activities' progress to assess their success against the expected goals. It focuses on the effectiveness and relevance of interventions, while it aims to provide information to all relevant stakeholders for planning future programmes. Evaluation is performed in specific points of time: (i) before activities take place, including targets' setting and system state estimation (ex-ante evaluation), (ii) during the programme while moving from one phase to another (in-itinere or mid-term evaluation), and (iii) after the activities conclusion (ex-post evaluation).

In respect of the programme implementation phases (Gaffey, 2013; Rogers, 2008; Wholey, 1983), monitoring tends to focus on the inputs, activities, and outputs, while evaluation concentrates mostly on the outcomes and impacts (Unicef, 1991), though this is not a clear cut.

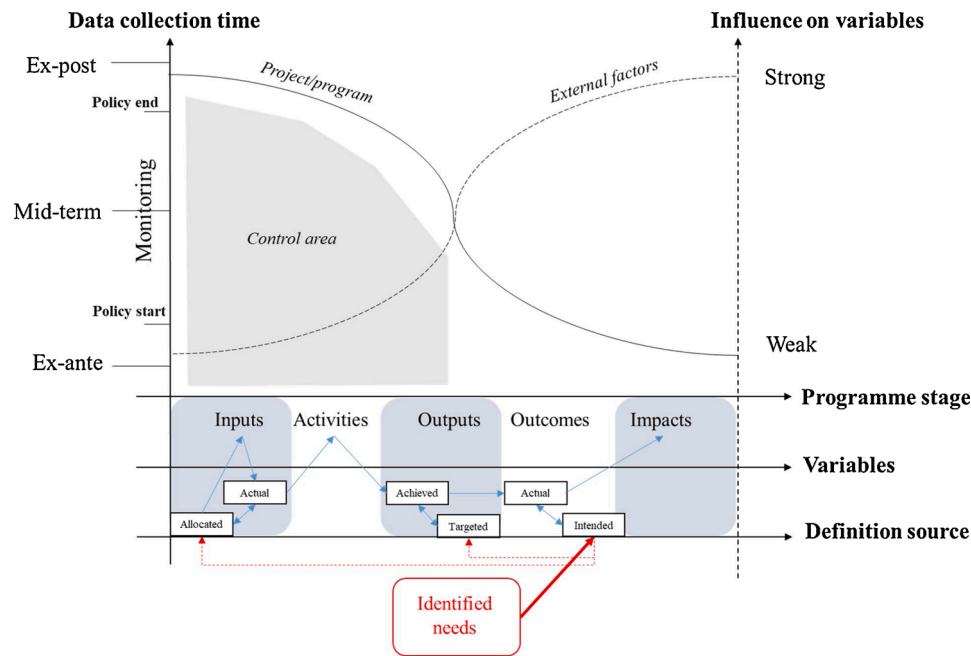


Fig. 1. Monitoring and evaluation integrated operative framework (Authors' own elaboration based on existing literature, adapted from Nigohosyan and Vutsova (2018)).

Inputs refer to the resources (e.g., human, economic, material) invested into a programme (Earl, Carden, & Smutylo, 2001; Rogers, 2008). Activities focus on the operations through which inputs are deployed to deliver specific outputs (Rogers, 2008; Smutylo, 2001), which are the immediate observable products resulting from programme activities (Earl et al., 2001; Lee et al., 2011). Outcomes (or results) constitute the short/medium-term consequences of an action's outputs that may associate to some extent to the implemented project (e.g., the direct effects that an intervention is meant to deliver) (Rogers, 2008). Outcomes are also seen as the interim benefits produced by outputs towards the impacts, which are meant as the external effects of a programme and ultimate or long-term results at a higher strategic level (Earl et al., 2001; Rogers, 2008).

Consequently, the influence of programmes is shifted progressively from inputs, activities, and outputs to outcomes and impacts. The explanation is attributed to: (i) the implementation agents that have different forms of control over the allocated and distributed resources, the implemented activities, and, in theory, the outputs produced (Nigohosyan & Vutsova, 2018; Smutylo, 2001), and (ii) the outcomes and impacts that are considerably influenced by external factors (Nigohosyan & Vutsova, 2018; Smutylo, 2001). The abovementioned concepts are integrated into a first-effort operative framework (Fig. 1).

2.2. Adapting the framework to Cohesion Policy

These concepts and frameworks have been further used to develop and analyse the CP monitoring and evaluation system (Barca & McCann, 2011; Gaffey, 2013; Nigohosyan & Vutsova, 2018). Financial resources (inputs) are allocated based on specific output targets deriving from the intended outcomes which, in turn, result from the identified needs (Barca & McCann, 2011). The actual spending of these resources supports the realisation of projects funded through CP (activities – e.g., airport expansion), which are expected to produce specifically targeted outputs (e.g., new terminals) (Barca & McCann, 2011). The direct immediate effects of the outputs' achievement constitute the outcomes (e.g., increased number of flights' arrivals/departures) (Nigohosyan & Vutsova, 2018). However, not only do the actual outcomes depend on the outputs, but also on other external factors (e.g., airlines' business decisions) (Barca & McCann, 2011). Finally, the impacts (e.g., increased

number of passengers/tourists) may partially depend on the policy effectiveness and, at a greater extent, on external factors outside the CP system (Barca & McCann, 2011; Nigohosyan & Vutsova, 2018), rendering the decoupling and assessment of the actual CP impact particularly challenging.

CP evaluation is principally related to outputs, results, and impacts, assessing the effectiveness and efficiency of the related programmes. It may occur (Bachtler & Wren, 2006; European Commission, 2018b; Lion & Martini, 2006): (i) during the planning phase (ex-ante), performed by LMAs and Member States, in the form of appraisal of the CP implementation context, indicators' selection, and baselines' and targets' development, (ii) during the programme execution (in-itinere and mid-term), conducted by independent regional and national evaluators with the scope of controlling whether the programme is on track and still relevant, and (iii) after the programme completion (ex-post), led by the European Commission in close cooperation with Member States and LMAs, involving an examination of the overall CP impact on different indicators to estimate programmes' effectiveness and enhance the design of the next programming periods. In this context, the evaluation of the implementation process (namely of its inputs and activities) "typically looks at how a programme is being implemented and managed", is attributed to the local authorities, and is "likely to be carried out in the early stages of implementation" (European Commission, 2018b, p. 28).

Instead, CP monitoring "observes programme implementation through, on the one hand, a continuous and systematic process of generating [...] information on implementation; on the other hand, through discussing these data sets in the Monitoring Committee"³ (European Commission, 2018b, p. 5). Therefore, "while primary responsibility for control rests with the audit and the [LMAs] in charge of each of the so-called Operational Programmes", Monitoring Committees are "an additional [...] mechanism for overseeing

³ "The monitoring committee are convened by the Managing Authorities responsible for each Operational Programme and are made up of officials from central and regional government, representatives from corporate and civic non-governmental organisations and, acting in an advisory capacity, officials from the European Commission. According to the relevant EU regulation, the role of the monitoring committees is to satisfy itself on 'the effectiveness and quality of the implementation of the Operational Programme[s]'" (Cartwright & Batory, 2012, p. 324).

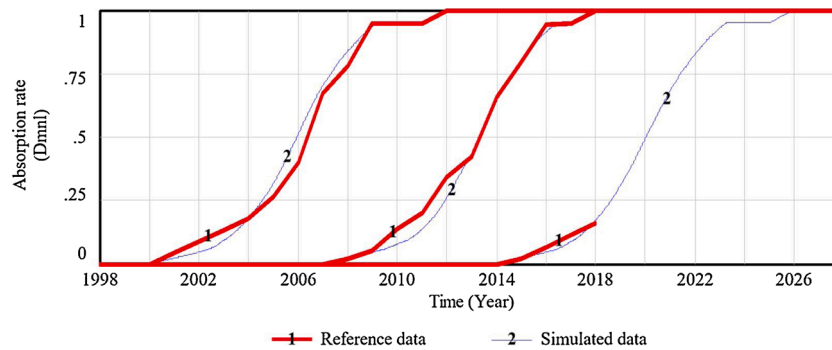


Fig. 2. Absorption rate calibration (thick red line – reference data; thin blue line – simulated data).

and advising on policy implementation” (Cartwright & Batory, 2012, p. 324). The monitoring scope is “to detect and quantify any deviation from initial plans and targets” (European Commission, 2018b, p. 5) for providing timely information to the management and enabling the correction of any potential source of issues. To minimise time-lags, the Commission further invites Member States, LMAs, and beneficiaries (or recipients) to upload the necessary data into the system throughout the year (e.g., on a quarterly frequency), and not just at the end of the year or the project. Programme monitoring is based on three categories of indicators: financial, output, and results (or outcomes) (European Commission, 2018b). Outputs are measured in physical or monetary units at the level of activities’ achievement. Results are defined in accordance to the specific programme goals, while data on the progresses and timing of collection are defined based on the specifics of the selected indicators.

Finally, “financial indicators relate to the total amount of eligible expenditure entered into the accounting system of the certifying authority and certified by that authority” (European Commission, 2018b, p. 7), are compulsory, and “may be used to monitor progress in terms of the payment of the funds available for any operation, measure or programme in relation to its eligible cost” (European Commission, 2018b, p. 7). These indicators are the closest instrument for inputs’ monitoring and refer basically to the absorption rate. However, the Commission acknowledges their limitations, though they are essential for monitoring programme performance, and allows LMAs to set additional indicators and targets exclusively for internal purposes (European Commission, 2018b). LMAs might (or probably do) use other internal dimensions for better controlling the inputs trend; however, there is no evidence in the literature nor public shared information. Therefore, in the regulation, the financial dimension (in the form of the absorption rate of expenditure) constitutes the only standardised, common, and publicly available proxy to monitor the inputs.

3. Methodological approach

3.1. System dynamics and modelling procedure

To evaluate the CP implementation system, we created an SD model to replicate the absorption rate of funds and to test whether this rate constitutes an efficient monitoring indicator of the inputs, as used by LMAs. SD can be considered as an appropriate method to tackle such problems since it focuses on complex systems, analysing longitudinal (i.e., over time) behaviours under policy interventions (Forrester, 1961; Sterman, 2000). From a technical perspective, an SD model constitutes a systemic structural map, including major cause-effect links and feedback mechanisms (Sterman, 2000). This operational conceptualisation allows to reconstruct the items and procedures constituting the system, thus providing a more practical understanding (Olaya, 2015). Arrows represent the links that connect a cause to its effect. The causal impact of each relationship is indicated with either a positive (i.e., both cause and

effect increase or decrease) or a negative (i.e., when cause increases, effect decreases and vice versa) polarity. The stock variables (presented as rectangles) express the states of the system due to the accumulation processes (i.e., mathematical integrations). The flow variables (presented as valves) refer to the rates that fill or empty the stock variables. In addition, a feedback loop is a circular sequence of causes and effects that is either balancing (i.e., self-stabilising) or reinforcing (i.e., self-increasing). If an initial increase in a variable leads to an eventual decrease (or increase) in the same variable, then the feedback loop is considered as balancing (or reinforcing) (or reinforcing). The formal model has been developed using the software Vensim®. To ensure transparency and replicability, the model and the related documentation are provided in the supplementary material (Appendix B), organised following the guidelines of Martinez-Moyano (2012) and Sterman and Rahmandad (2012).

The model structure was presented, developed, revised, and validated in: (i) two workshops, organised grounding on a typical group model building approach (Vennix, 1996), in the context of a Horizon 2020 project attended by CP experts (i.e., EU policy-makers, regional officers, researchers), and (ii) twelve interviews, conducted in a disconfirmatory approach (Andersen et al., 2012), with regional officers, researchers, and CEOs. In the Appendix A, details about these meetings are reported. Participants and interviewees were given an enlarged printed copy of the model and invited to modify its structure on paper. This procedure was used as a starting point for discussion and acted as a boundary object to structure the interactions (Black, 2013). Each meeting was recorded, if permitted by interviewees and participants, and conducted by at least two members of the research team to allow for an adequate flow and notes’ collection. Then, participants’ and interviewees’ notes on the model, recordings, researchers’ notes, and impressions were compared to obtain a consistent interpretation of elicited mental models (Doyle & Ford, 1999). Overall, the presentation of updated model versions and their discussion during workshops and interviews has acted as a continuous validation process of the conceptual structure of our work.

Rigorous validation tests have been performed to increase the confidence in the model and its outputs, namely: structure and parameter confirmation test, dimensional consistency, formal inspections, walk-throughs, extreme conditions tests, behaviour sensitivity tests, and modified-behaviour predictions (Barlas, 1996; Sterman, 2000). In addition, to verify whether the structure is reliable, the model has been calibrated to replicate total absorption rate (i.e., of both EU and regional shares of contribution) of the European Regional Development Fund (ERDF), as a component of the structural funds, over three policy cycles 2000–2006, 2007–2013, and 2014–2020 in the region of Emilia Romagna in Italy. Although several LMA officers around the EU have participated in the workshops and interviews, Emilia Romagna was selected as a case study since the respective LMA was a partnering organisation of the Horizon 2020 project, as well as due to the availability of regional absorption data. More specifically, the EU share data

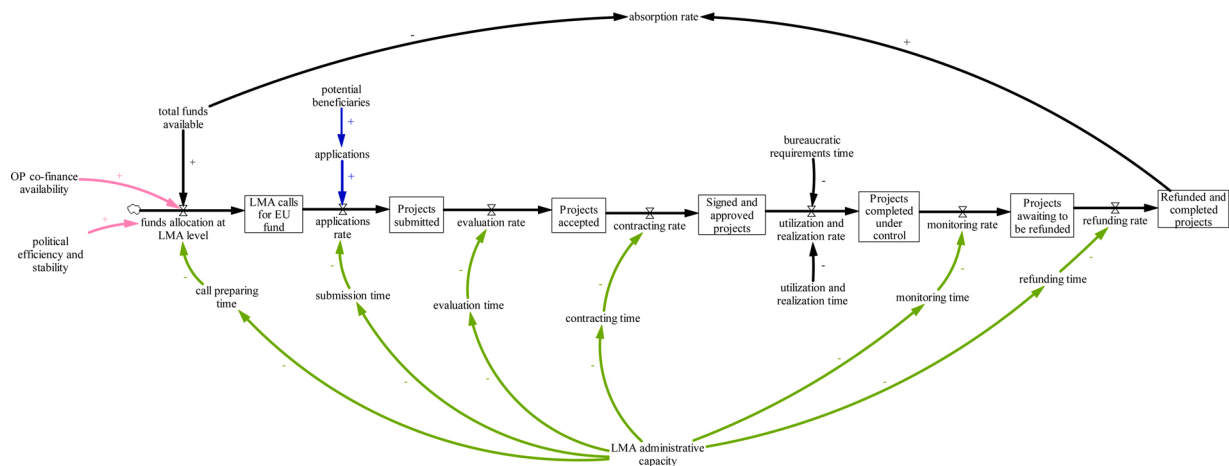


Fig. 3. Systemic map of Cohesion Policy implementation 'pipeline'.

were retrieved from official internally-communicated annual reports via e-mailing (European Commission, 2003–2018). For the 2000–2006 cycle, EU data availability starts in 2003 (for the previous years, a constant funds' expenditure is assumed). For the 2014–2020 cycle, the data were only partially available till 2018. Regarding the regional share, the data were obtained from the OpenCoesion database⁴, which is the Italian initiative of centralised collection and sharing of CP data.

Fig. 2 illustrates the manner in which the developed model replicates the real data available (i.e., reference mode). The model outputs replicate the reference modes' S-shaped growth to a sufficient extent, further increasing the confidence in the model structure's ability to capture the complex dynamics of CP implementation at an LMA level. In fact, as the system slowly starts to absorb the new cycle's funds (due to the initial delays involved), the absorption rate steepness increases. Towards the end of the cycle, absorption slows down due to the 'natural' saturation effect (there are fewer funds and therefore the absorption decelerates) till it reaches approximately 95 % (when LMAs have to wait the EU and the delays involved to process the final payments before providing the last reimbursement). The ability to replicate reality represents an important target of the SD modelling approach, since achieving this goal may provide a first formal validation of the quality of the model (Barlas, 1996; Forrester & Senge, 1980; Sterman, 2000).

3.2. Model structure: Cohesion Policy implementation map

The system structure reflects the major CP implementation processes, including the main flow of structural funds from the EU through the regions to the beneficiaries, as well as all parameters and factors affecting this procedure. Since the analysis includes three key CP players, the model attempts to depict the multi-level nature of the scheme (Hooghe, 1996; Piattoni & Polverari, 2016) into an operational framework. Thus, in this case, SD allows to deconstruct the elements that constitute the CP implementation for understanding the causes of specific funds' absorption patterns over time. The model is portrayed in an aggregated qualitative form (Fig. 3) to facilitate the readers' understanding. The system map reports the underlying logic of the main drivers that affect CP implementation in the model; specifically, the main flow of funds (i.e., pipeline) is presented in black, while the drivers that affect the implementation processes refer to the contextual external factors (pink arrows), the beneficiaries' demand (blue arrows), and the LMA administrative capacity (green arrows).

Regarding the funding pipeline, the "total funds available" by the EU increase the allocation rate among the regions ("funds allocation at LMA

level"), which determines the amount of "LMA calls for EU funds". The more funds are allocated, the more calls are prepared by the LMAs. "Potential beneficiaries" who apply to the calls ("applications") are accounted in the "applications rate" mechanism; depending on the number of calls, potential beneficiaries submit their application, accumulating into the stock of "Projects submitted". The submitted applications are evaluated ("evaluation rate") and all evaluated projects accumulate in the "Projects accepted" stock. These projects are then signed ("contracting rate") and the approved ones are transferred to the "Signed and approved projects" stock. Subsequently, the projects are put into action ("utilization and realization rate") and after the bureaucratic requirements' control, projects are completed and accumulate in "Projects completed under control". In this stage, they are assessed ("monitoring rate") and those passing the assessment process move to the state of "Projects awaiting to be refunded". After technicalities are solved and processed, beneficiaries are finally refunded ("refunding rate") and all projects accumulate in the "Refunded and completed projects" stock. The ratio of the funds actually spend in the refunded projects to the total funds available constitutes the funds' "absorption rate". Concerning the external contextual factors, the "OP co-finance availability" (OP stands for Operational Programme), which refers to the LMA's economic contribution to the CP, and the "political efficiency and stability" positively affect the regional allocation of the funds. From the beneficiaries' part, the more "potential beneficiaries" are interested, the more projects applications are submitted. Finally, an increase in the "LMA administrative capacity" can reduce the time needed to perform each process step of the main pipeline by eliminating systemic delays. Finally, shorter process times increase the related rates, rendering the pipeline flow quicker and leading to more efficient processes and more projects completed.

4. Indicators' dynamic behaviour

4.1. Dynamic adequacy test of absorption rate

Three different scenarios are simulated to test the adequacy of the absorption rate as a monitoring indicator of the inputs implementation. To perform these tests, the model is simulated for the upcoming policy cycle by keeping constant the previous cycles assumptions and conditions. Exogenous extreme shocks are simulated to assess the receptivity of the absorption rate indicator to the exceptional disturbances. This extreme cases' approach highlights the shortcomings of the financial indicator. There are three cases and they constitute the 'entry points' of drivers into the implementation pipeline (i.e., beneficiaries' demand, external factors, LMA administrative capacity) which can disturb the regular funds' expenditure.

⁴ <https://opencoesione.gov.it/en/>

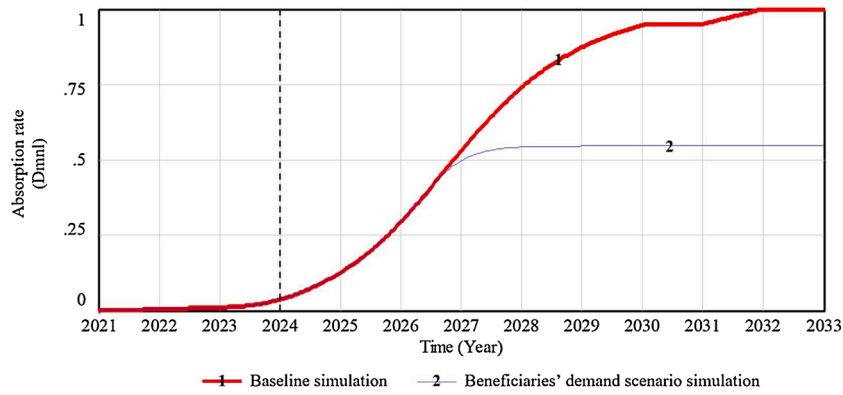


Fig. 4. Absorption rate (thick red line – baseline simulation; thin blue line – scenario simulation) – Beneficiaries' demand scenario (zero applications from 2024).

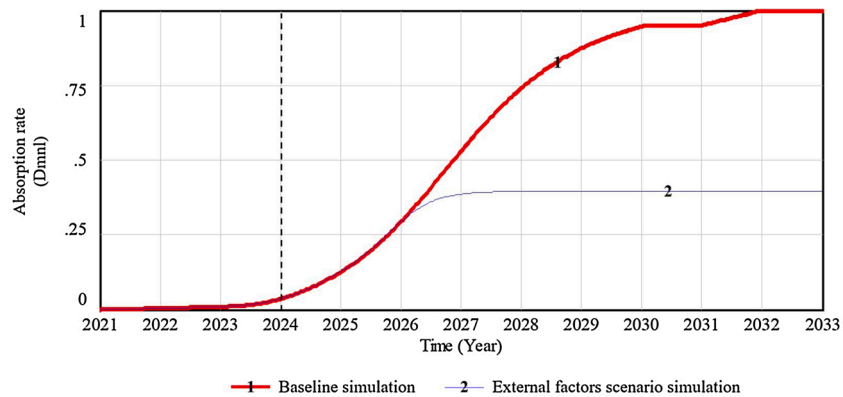


Fig. 5. Absorption rate (thick red line – baseline simulation; thin blue line – scenario simulation) – External factors scenario (zero calls from 2024).

Concerning the beneficiaries' demand, a scenario in which in the beginning of 2024 no potential beneficiary applies for funds is simulated (Fig. 4). In this case, the flow through the pipeline is disrupted since the causal link between applications and the pipeline is interrupted. The thick red line represents the base run in which absorption occurs regularly without any issues, while the thin blue one shows the absorption behaviour in case from 2024 no applications are submitted anymore and suddenly demand goes to zero. As shown, the absorption regularly increases, as if nothing has happened, until almost 2027 and only after this date the absorption curve flattens. Therefore, the problem manifests itself in the absorption behaviour with approximately three years of delay (all estimations are performed qualitatively through observation by comparing the curves' heights, slopes, and patterns (Sterman, 2000),

which is deemed to be a sufficient approach given the model's purpose (Groesser & Schwaninger, 2012)). These delays are involved in the implementation pipeline; while demand for funds occurs at the beginning of the process, the calculation of the absorption is performed at the end. Thus, if a problem occurs at the beginning of the pipeline, the rest of the projects accumulated in the pipeline are carried on and the problem will emerge only after the pipeline will be emptied.

The second shock scenario refers to external factors and simulates the case in which, at the beginning of 2024, the LMA stops releasing CP calls. This case is simulated by increasing to the maximum the political instability in the model (this shock could also be generated by a lack of co-finance availability), which may not provide the necessary legal and political direction to the LMA to prepare and publish the calls (Fig. 5).

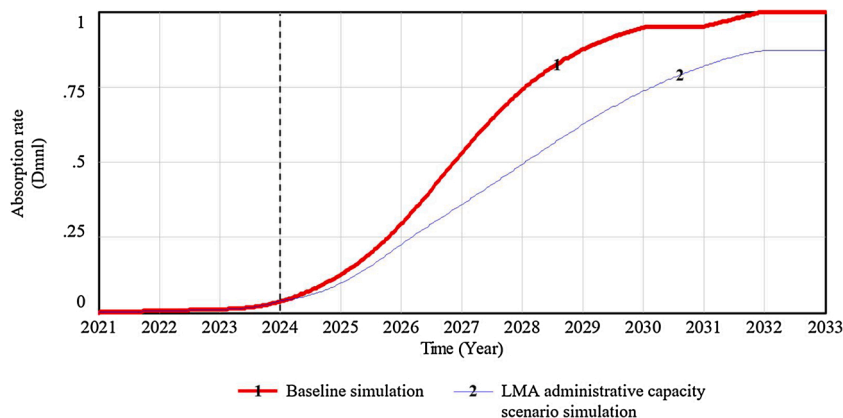


Fig. 6. Absorption rate (thick red line – baseline simulation; thin blue line – scenario simulation) – LMA administrative capacity scenario (60 % decrease in LMA's staff from 2024).

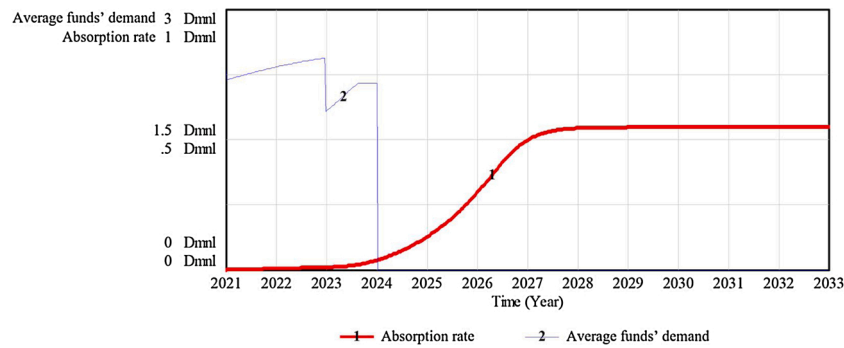


Fig. 7. Absorption rate (thick red line) vs. average funds' demand (thin blue line) – Beneficiaries' demand scenario (zero applications from 2024).

Once again, it should be noted that the absorption rate reports the shock in the behaviour only after 2026. This latter is visible slightly earlier compared to the previous case since, by closing the pipeline earlier in the stream, a lower amount of projects is introduced into the processing. This fact amplifies the shock in the implementation system.

Finally, the third case is associated with the LMA administrative capacity. Specifically, at the beginning of 2024, there is a 60 % decrease in the LMA's staff which is not counterbalanced by any hiring (Fig. 6). In respect of the modelling procedure, this scenario is translated in a remarkable reduction in the LMA administrative capacity and, subsequently, in a longer time to process the applications and the projects through the pipeline. The simulation's results indicate that the absorption rate takes a different path from the baseline simulation after a few months. Although the two behaviour types become separated relatively shortly, the distance between the two curves is rather evident only after a few years. This observation might also mislead policy-makers since they may underestimate the importance of absorption discrepancy (e.g., they might consider it as a normal statistical fluctuation).

4.2. Towards a new set of original operational indicators

Given the alarming delay with which the absorption rate captures the system's extreme conditions, we develop and test a new set of operational indicators which could support the timely detection of variations in the system's behaviour. The underlying idea is that these new operational indicators will complement financial proxies (e.g., absorption rate) during the monitoring of performance, allowing Commission, Member States, and LMAs to know in time where problems are located and, thus, solve them quickly. In addition, the indicators' public availability would improve the understanding of CP implementation among researchers, stakeholders, journalists, and the society in general. Therefore, a set of three quantitative indicators, identified based on the main systemic drivers that affect the CP implementation upstream in the pipeline (beneficiaries' demand, external factors, and LMA administrative capacity), are outlined and simulated. These indicators aim to monitor the 'strength' and the impact of these drivers. They have been developed following the EU principles for: clarity, unambiguity, easiness to understand, closely linked to the activities, with transparent units of measurement, and allowing for periodic measurements (Barca & McCann, 2011; European Commission, 2018b).

The first indicator, named as *funds' demand*, has the ability to monitor the LMA calls for projects that remain unanswered. More specifically, it assesses the success of a call among the local community, hence it provides an indication of the number of potential beneficiaries who apply for a specific call. It is calculated by dividing the amount of funds that the potential beneficiaries request by the total funds' available for a certain call, providing a ratio of call fulfilment:

$$funds' demand = \frac{fund\ demanded\ by\ potential\ beneficiaries}{funds\ available}$$

The indicator's value can be explained as follows. If the *funds' demand* is lower than one, the calls have been partially gone unanswered. If it is zero, nobody applied. The closer it gets to 1, the more the request for funds through the submitted applications reach the whole allocated amount. In particular, if the indicator equals one, the demand perfectly matches the supply. In other words, the request for funds equals the allocated sum. Finally, if it is greater than one, there are more applications than the number of places available meaning that the request for funds exceeds the allocated sum. The last scenario is the usual one in LMAs with good performances. Receiving more project proposals than needed allows the LMA to have a considerable amount of projects to assess in case some do not pass the evaluation phase (e.g., due to low quality), thus avoiding the call to go partially unfulfilled. Paradoxically, some LMAs pointed out that a value of *funds' demand* equal to more than one can have a discouraging effect for the beneficiaries; if applicants have their project proposals rejected too often, they might be discouraged from submitting again. Thus, an LMA might also decide to use other regional and/or national funding sources to top up the call and fund additional projects in case they fulfil the quality standards.

However, even high-performance LMAs can draft a lowly appreciated call, while low-performance LMAs a highly appreciated one. Thus, the *average funds' demand* among several calls prepared can be more valuable from a systemic perspective. In this case, the indicator is developed by summing all call performances and then dividing them by the number of calls for a certain period:

$$average\ funds' demand = \frac{\sum(funds' demand)}{number\ of\ calls\ in\ a\ certain\ period}$$

This indicator can be useful to understand whether absorption issues arise from the demand (beneficiaries) or the supply (LMAs) side. Namely, in the case of a CP programme with a bad performance, the system's behaviour should be investigated in terms of beneficiaries applications; if the beneficiaries do not apply, the problem is probably located on the demand side. In this case, the only feasible policy intervention at the moment seems to be the engagement and discussion with the local community to analyse the reasons behind the low number of applications (e.g., lack of proper information, lack of beneficiaries' staff capability in preparing applications, costly applications, lack of credit to co-finance), allowing the LMAs to remedy the undesired situation. Overall, not only does this indicator assist in identifying in which stage of the CP process low performance arises but it also highlights when performance starts to be undermined before it is identifiable in the absorption rate. Although this indicator is based on monetary values, such as absorption rate, it has an operational nature that provides a different dimension. In fact, this indicator allows LMAs to intervene earlier and avoid problematic situations. In Fig. 7, a comparison between the absorption rate and the *average funds' demand* is proposed in the initial scenario of a halt in the beneficiaries' applications from 2024 onwards. As shown, the *average funds' demand* indicator immediately collapses. In Fig. 8, the comparison of the *average funds' demand* between the baseline

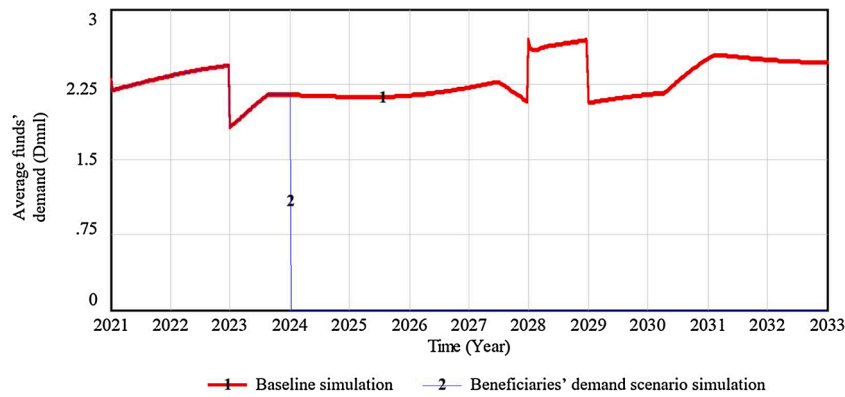


Fig. 8. Average funds' demand (thick red line – baseline simulation; thin blue line – scenario simulation) – Beneficiaries' demand scenario (zero applications from 2024).

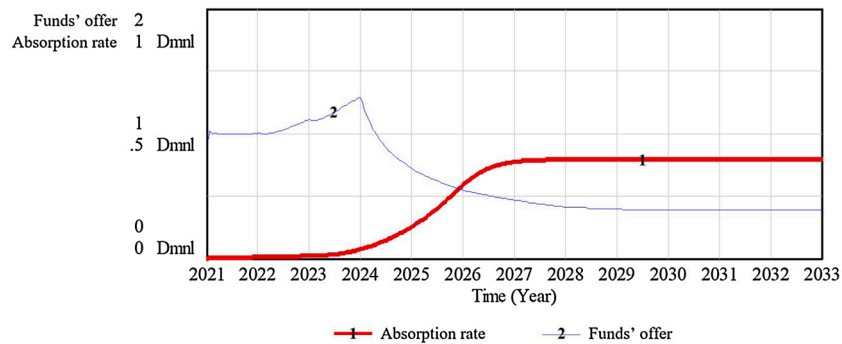


Fig. 9. Absorption rate (thick red line) vs. funds' offer (thin blue line) – External factors scenario (zero calls from 2024).

simulation and the scenario of zero applications is depicted.

A variation in this indicator could refer to the number of projects accepted (i.e., the applications which meet the eligibility criteria and pass the initial selection). This indicator can be named as **actual average funds' demand**, and, compared to the *average funds' demand*, it further indicates the quality the projects submitted. If there is a considerable discrepancy between the two, then the quality of the applications is low, since there is a high number of projects rejected that makes the actual average ratio lower than the average ratio.

If the beneficiaries apply but a problematic absorption behaviour still occurs, dysfunctions may be located on the LMA side. To explore them better, two additional indicators have been developed for supporting the systemic search for this behaviour. The CP **funds' offer** represents the total amount of the funds committed in all different phases of the implementation process in a specific point in time compared to a

reference desired value (e.g., a baseline value defined by the Commission or the LMA in ex-ante). Specifically, the funds committed comprises all funds that are allocated in calls, committed to projects under evaluation, approval, signature, execution, monitoring, and refunding queue, as well as the already refunded funds. The proposed indicator is defined as follows:

$$funds' offer = \frac{funds\ committed}{desired\ funds\ committed}$$

To some extent, this indicator can be considered as similar to the absorption rate since it focuses on the monetary value that an LMA should spend. Nonetheless, the *funds' offer* is a much broader concept and allows for the monitoring of the LMA effort in real-time, since the day one of the policy cycle. In fact, the absorption rate cannot be calculated from the beginning of each cycle, since no absorption is made

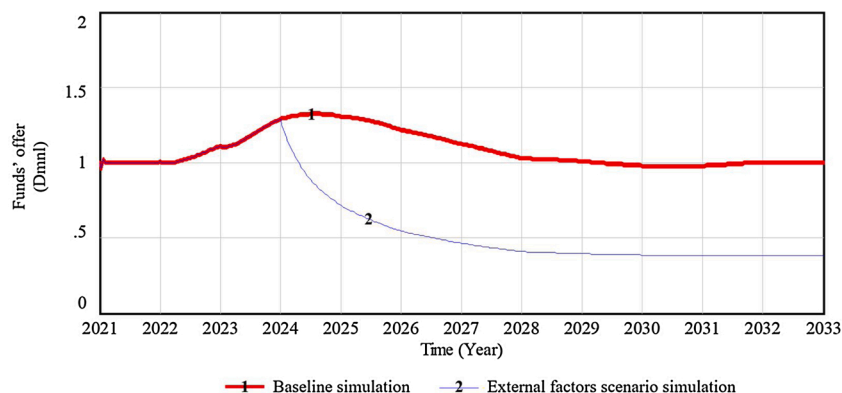


Fig. 10. Funds' offer (thick red line – baseline simulation; thin blue line – scenario simulation) – External factors scenario (zero calls from 2024).

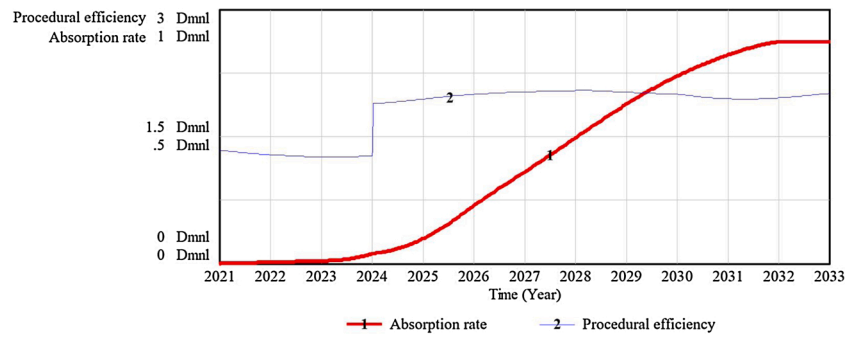


Fig. 11. Absorption rate (thick red line) vs. procedural efficiency (thin blue line) – LMA administrative capacity scenario (60 % decrease in LMA’s staff from 2024).

(it takes time to complete the whole pipeline of processes till the expenditure certification and refund). Therefore, the *funds’ offer* reports which are the actual progresses of the LMA at any time point and goes beyond the absorption rate or the certified expenditure. However, this indicator works best if combined with the *average funds’ demand* and the absorption rate. If, while analysing an LMA with a low performance, the *average funds’ demand* is sufficiently high, then attention should be paid at the *funds’ offer*; the latter might show that although the calls are fulfilled, the amount of money allocated, committed, and refunded is rather low compared to what it should be in order to reach an adequate absorption. Fig. 9 reports a comparison between the absorption rate and *funds’ offer* reaction to the scenario in which from 2024 no new funds are allocated to any call due to local political instability. The simulation results showcase that the proposed indicator exhibits almost immediately the effect of the external environment on CP implementation, allowing for timely corrective interventions by the Member States and the Commission (*funds’ offer* goes above one in the beginning of the period, indicating that the LMA has a better performance than the defined baseline). Fig. 10, instead, compares the values of *funds’ offer* between the baseline simulation in which the implementation proceeds regularly and the scenario of zero funds’ commitment in calls.

Once again, this indicator could allow for the understanding of the interventions needed by the LMA. However, if there is a sufficient *funds’ offer* with a proper *average funds’ demand* and the absorption rate is still too low, the problem might be located in the length of the different procedural steps that build up the main CP pipeline. Therefore, a third indicator needs to be acknowledged as a time performance indicator. Specifically, monitoring the time needed to perform the necessary procedures can indicate how quick the LMA is in performing its duties. In fact, this indicator aims to offer a comparison between the reference average time that each step requires to be performed (i.e., the time that the European Commission considers as a standard time for each step)

and the average time that the LMA needs to perform each step in reality. Examples of time steps may refer to the time needed to evaluate the projects’ applications received for a call or the time needed to sign the contracts with the beneficiaries, as indicated in the system map. The reference and actual times are finally summed for all steps and the **procedural efficiency** is calculated as the ratio of the respective sums:

$$procedural\ efficiency = \frac{\sum_1^n (Step\ n\ time)}{\sum_1^n (Step\ n\ reference\ time)}$$

If the time performance indicator is greater than one, this fact indicates that the LMA needs more time to complete the steps on average, while if it is less than one the LMA is quicker than expected. If it is one, the LMA requires, on average, the reference time as defined by the Commission. This indicator could provide real-time information on the CP implementation procedural performance. In addition, its deconstruction on step level could showcase where the bottlenecks are located. Fig. 11 compares the behaviour of the absorption rate and the *procedural efficiency* indicator in the scenario in which the 60 % of the LMA staff devoted to CP implementation is fired from 2024 onwards. As indicated, the *procedural efficiency* manifests this problem immediately. In addition, Fig. 12 portrays the comparison between the baseline simulation of regular implementation and the extreme scenario of staff decrease. The baseline curve stands around 1.2–1.3 meaning that, in the base scenario, the procedure takes 20–30 % more than the optimal time hypothesised. The two curves follow different paths in 2024 due to the change in the conditions; the curve simulating the extreme case stands above the base one, approximately around 2–2.2, meaning that it takes more than double time compared to the optimal one.

If it was possible to know the real-time that an LMAs needs to perform its duties at each phase of the project implementation, then

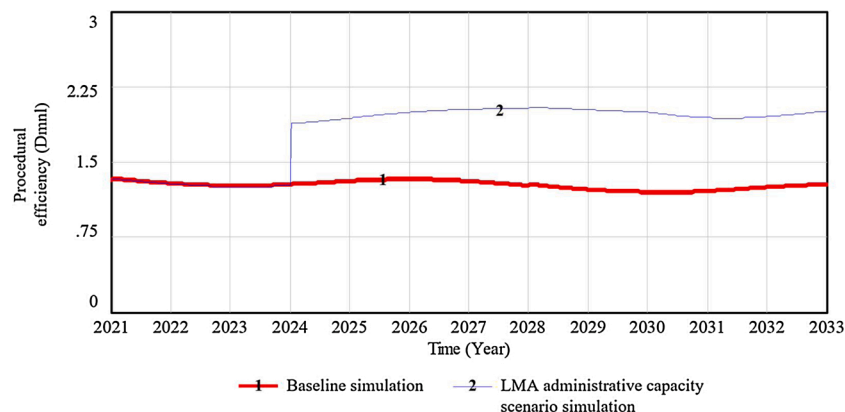


Fig. 12. Procedural efficiency (thick red line – baseline simulation; thin blue line –scenario simulation) – LMA administrative capacity scenario (60 % decrease in LMA’s staff from 2024).

these data could provide further information about noteworthy aspects. For example, for similar project types, if the realisation phase (in which the LMA performance is not the main determinant since this phase refers to the beneficiaries implementing the project) is monitored to be much longer than the average one, the fault could be assigned to the ‘gold plating’ (European Court of Auditors, 2016; Gandolfo, 2014). In particular, this concept refers to the presence of redundant national and regional bureaucratic procedures in addition to the CP regulations, which ultimately end up increasing the administrative workload for LMAs and beneficiaries.

5. Analysis and discussion

5.1. Outcomes’ analysis

If financial proxies (i.e., absorption rate) grasp with a considerable delay extreme system conditions, they will detect minor implementation issues with even more difficulty. A low absorption rate might arise due to multiple reasons (e.g., low beneficiaries’ demand, political instability, low LMA capacity); however, identifying the source and the locus of undesired dynamics in the system by only exploring the absorption rate may be quite challenging. Thus, deconstructing the problem to its root causes and intervening for solving it emerge as crucial. In addition, the absorption rate cannot showcase the use of debatable strategies (e.g., retrospective projects’ use, regional co-finance reduction) by LMAs with expenditure difficulties, which increase the numerical value of the rate, yet without increasing the impact of the CP intervention. Moreover, the absorption rate cannot provide a timely assessment of the inputs’ implementation due to delayed detection of issues in the implementation system. In fact, as absorption refers to downstream variables, it indicates if there are problems upstream in the CP implementation system only at the end of the pipeline flow given the intrinsic delay involved. For example, we may consider an LMA with low performance due to unfulfilled calls; by using the absorption rate, this problematic behaviour will be discovered only years later when the indicator manifests the undesired trend. For instance, in the initial years, it is challenging to realise whether there are performance issues since, in this period, the absorption rate is generally low and is affected by pre-finance mechanisms⁵. Consequently, this delay in perception entails, in turn, a delay in the corrective interventions that could lead the system towards a more desirable state. Given these specific characteristics, the absorption rate is a rather uninformative indicator for monitoring and, thus, it cannot provide adequate support to gain comprehensive understanding of the underlying phenomena behind the implementation process.

5.2. Policy implications and opportunities

In this context, a set of three upstream indicators are developed to be used by the different authorities of the multi-level governance scheme (e.g., European Commission, Member States, LMAs), since the necessary data can be easily collected and shared by the LMAs. The proposed indicators aim to be quite uncomplicated to allow for easier understanding and, thus, match the cognitive needs (Barca & McCann, 2011; Lion, Martini, & Volpi, 2004) of stakeholders, policy-makers, practitioners, and citizens. Given that the EU currently demands only financial proxies for inputs monitoring and gives quite some autonomy to the LMAs, an enhanced standardised set of monitoring indicators could create a systemic dashboard of real-time dynamic monitoring. This analysis could increase the effectiveness of monitoring as a collection of useful data for improving performance and understanding, as well as guarantee that

⁵ The European Union advances a specific amount of CP funds to the local authorities as soon as the CP programmes are approved (European Parliament & the Council of European Union, 2013)

LMAs are approaching monitoring in the best-known way. On the one hand, the fact that regions encounter performance problems could be due to their lack of proper monitoring tools and this could be prevented with the new systemic monitoring dashboard. On the other hand, the proposed approach would better fulfil the summative and formative scopes of the monitoring practice (Wholey, 1996) compared to the use of absorption rate as the only available indicator.

Recently, the Commission has suggested shifting the focus of CP implementation from financial performance and compliance to performance assessment towards tangible results (European Commission, 2019). Financial dimensions have monopolised the CP implementation so far and have generated a wide range of issues (Polverari, 2015; Rainoldi, 2010) among which the misuse of absorption rate as a proxy for monitoring performance. Specifically, the Commission proposes “payments from the Commission to the Member State or region conditional on the achievement of pre-agreed results/outputs or completion of policy actions or processes” (European Commission, 2019, p. 6). In this radical and historical potential change in the CP system, inputs’ monitoring could also be updated with a more structured, standardised all over Europe, dynamic, systemic, and operational approach. Moreover, if deemed as valid, this approach could ‘spill over’ into the Member States monitoring and evaluation culture, as it happened in the past (Viñas, 2009).

5.3. Limitations and challenges

The implementation of the proposed research still requires further refinements before being used in practice. The developed SD model is a representation of the CP implementation system, constituting a conceptual simplification. Therefore, the proposed indicators may not entirely grasp the complexity that emerges during the structural funds’ allocation and expenditure and thus they cannot be immediately used by the LMAs. The systemic indicators can be interpreted as a preliminary and necessary directional advancement that should be followed by rigorous adaptation and definition based on the day-to-day LMA and EU procedures.

The indicators’ revision should be based on available data, which can be potentially collected by the LMAs, as well as on how these eventual requirements posed by the new indicators can be integrated into the system of data registration. Direct experience of some Member States shows that they have been already collecting types of data close to what is needed and based on what is proposed (OECD, 2015). For instance, OpenCoesione, the Italian initiative developing a database of CP data throughout the country, collects a variety of information about the implementation process from LMAs and renders them publicly available. Then, the collected data could be provided in a transparent manner for all EU countries in the CohesionData⁶ web space to increase accountability and political responsibility.

LMAs probably collect information which is partially suitable for developing the suggested indicators. Although using the new indicators might increase complexity instead of simplifying it, the calculation is quite simple and may alleviate the burden from the LMAs that should otherwise develop assessment techniques by themselves. Similarly, the data collection could increase the work load of the LMA staff; however, given the increased automation level nowadays, a minor increase is expected. In general, the LMA workload and specific needs have to be carefully considered (since LMAs might be sensitive to task variations in this sense) in case policy-makers aim to diffuse the proposed approach (Masana & Fernández, 2019). Nevertheless, the benefits of timely intervention, transparency, and more accurate accountability could make the case of outweighing the potential costs.

Moreover, the definition of the reference values for the ideal funds’ offer over time and the procedural timings might be rather challenging.

⁶ <https://cohesiondata.ec.europa.eu/>

In fact, the problem of setting benchmarks for CP implementation has been highlighted in the literature (Marchante & Ortega, 2010). However, optimal performances from previous policy cycles could be used as a starting point for baseline definitions. Finally, the proposed indicators need to be handled carefully; it is crucial to avoid any potential goal displacements (i.e., used with other meanings and scopes instead than the original ones) (Perrin, 2011), which roughly happened to the absorption rate. In addition, the new indicators might tend to reduce participation and monopolise the decision-making process, especially if they are attributed with a broader meaning than they have (e.g., used as proxies of LMA quality of governance).

6. Conclusions

6.1. Major insights

This paper analyses the adequacy of the most used financial indicator for inputs monitoring in the context of CP, namely the absorption rate of structural funds. To that end, a SD model replicating the expenditure process of the ERDF funds has been built to simulate different extreme scenarios and observe whether the absorption rate provides timely and precise information about the state of the CP implementation system. The results highlight that, in three different cases in which a sudden change occurs (i.e., low beneficiaries' demand, lack of funds allocation due to political instability, delayed implementation process due to low LMA capacity), the absorption rate captures only with a notable delay the new conditions. Therefore, we claim that the absorption rate could delay and mislead real-time monitoring and, thus, any eventual corrective interventions, as it assesses the performance downstream in the implementation pipeline. In this respect, the proposed systemic structure allows for the development of three original operational indicators based on the different driving forces of the implementation process (i.e., average funds' demand, funds' offer, procedural efficiency), which are anticipated to offer a timely snapshot on the system state and complement the absorption rate.

Overall, this paper proposes a new systemic and operative approach to the monitoring of CP inputs, beyond the financially oriented approach based on the use of absorption rate. The term '*systemic*' refers to the fact that monitoring can benefit from holistic understanding of the implementation structure and its state, while the term '*operative*' emphasises the importance of practical aspects (Saunders, 2011) in monitoring. In this respect, the attention towards the integration of the SD concept within the CP is gaining attention in the fields of implementation (Kvaca & Kokeš, 2018) and analysis of the CP processes, procedures, and organisations (Cunico et al., 2020; Smeriglio et al., 2015). In respect of monitoring and evaluation, there are no evident theoretical obstacles to such integration and it could potentially occur as it happened in other domains (e.g., sustainability indicators, Meadows, 1998). In general, this study can provide an additional example of the benefits of using principles of systems' theories for evaluation and monitoring purposes, further nurturing the discussion on this integration (Gates, 2016, 2017; Hassmiller Lich et al., 2017; Hummelbrunner, 2011; Midgley et al., 2008). In addition, the new possible approach about payments proposed by the Commission, in which financing is not linked to financial absorption but to the achievement of tangible results of the LMAs (European Commission, 2019) shifting away from the money-oriented compliance (Polverari, 2015), could be a fertile political ground to embrace new monitoring frameworks. This new approach could lead to better attribution of political accountability of the LMA concerning CP implementation performance (Davies & Poverari, 2011; Gore & Wells,

2009; Poverari, 2015). This is still a critical issue in several Member States (Marra, 2018; Poverari, 2015) since the correct attribution of the responsibility for performance in CP implementation can shape the citizens' perception of local, national, and European institutions.

6.2. Research and policy directions

This study highlights the importance of the expansion of knowledge during the CP monitoring phase, giving more attention to the topic by generating guidelines for practitioners. In respect of the proposed indicators, additional work should be directed towards their refinement, adaptation, and harmonisation against the numerous nuances of the implementation process; the next step should render these conceptual ideas ready to be put in action. Future research efforts could perform more extensive experimentation of simulation tools, especially of operational research methods such as SD, for studying the CP monitoring and evaluation system deeply.

In conclusion, this study is expected to raise awareness of the limited adequacy of the absorption rate as a monitoring indicator for the CP inputs. This might have been already detected by some LMAs that adopt similar indicators as the proposed ones. However, up to now, there is no transparent, robust, and shared information on the manner in which regions proceed with performance monitoring. Therefore, the EU should provide additional details on the monitoring side of the policy, standardise this procedure to the best possible extent, and promote a transparent sharing of the data collected. Finally, a more accurate, timely, and transparent monitoring of the inputs could decrease the number of regions in which absorption issues occur, since corrective interventions could take place in time. In addition, this inputs' analysis could feed practitioners and researchers with more precise data regarding the CP implementation process.

Author statement

Conceptualization: G. Cunico; E. Aivazidou; E. Mollona

Data curation: G. Cunico; E. Aivazidou

Formal analysis: G. Cunico; E. Aivazidou

Funding acquisition: E. Mollona

Investigation: G. Cunico; E. Aivazidou

Methodology: G. Cunico; E. Aivazidou; E. Mollona

Project administration: G. Cunico; E. Aivazidou

Supervision: E. Mollona

Validation: G. Cunico

Roles/Writing - original draft: G. Cunico; E. Aivazidou; E. Mollona

Writing - review & editing: G. Cunico; E. Aivazidou; E. Mollona

Declaration of competing interest

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Acknowledgement of funding

This research has been conducted in the context of the project: "Perception and Evaluation of Regional and Cohesion Policies by Europeans and Identification with the Values of Europe" (project's acronym: PERCEIVE) that received funding from the European Community's Horizon 2020 Programme (H2020-REFLECTIVE-2014-2015) under the Grant Agreement No. 693529.

Appendix A

Table A1

Table A1
Interviews and workshop information.

Type	Number of participants	Participants' job titles	Organisation type	Date and length
Interview	1	CEO	Private company (both beneficiary and intermediary company)	17 th April 2018 105 minutes
	2	Researchers	University	11 th June 2018 100 minutes
	1	Journalist	Newspaper	25 th June 2018 140 minutes
	1	Officer	LMA	10 th July 2018 150 minutes
	1	Employee	European Structural Fund (ESF) funded project participant	17 th September 2018 50 minutes
	1	Officer	Europe Direct	20 th September 2018 110 minutes
	1	Researcher	University	25 th September 2018 110 minutes
	1	Officer	LMA	25 th September 2018 80 minutes
	1	CEO	Private company delivering ESF-funded courses	5 th October 2018 60 minutes
	2	Researchers	University	5 th October 2018 70 minutes
	2	Officers	OpenCoesione (Italian Open Government initiative about Cohesion Policy)	2 nd December 2018 100 minutes
	2	Officers	LMA	26 th February 2019 150 minutes
Workshop	approx. 20	LMA officers, EU officers, researchers		26 th October 2018 300 minutes
Workshop	approx. 15	LMA officers, EU officers, researchers		18 th June 2019 120 minutes

Appendix B. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.evalprogplan.2021.101964>.

References

- Aivazidou, E., Cunico, C., & Mollona, E. (2020). Beyond the EU structural funds' absorption rate: How do regions really perform? *Economies*, 8(3), Article 55. <https://doi.org/10.3390/economies8030055>
- Andersen, D. L., et al. (2012). The disconfirmatory interview as a strategy for the assessment of system dynamics models. *System Dynamics Review*, 28(3), 255–275. <https://doi.org/10.1002/sdr.1479>
- Bachtler, J., & Ferry, M. (2015). Conditionality and the performance of European structural funds: A principal–Agent analysis of control mechanisms in European Union Cohesion Policy. *Regional Studies*, 49(8), 1258–1273. <https://doi.org/10.1080/00343404.2013.821572>
- Bachtler, J., & Michie, R. (1995). A new era in EU regional policy evaluation? The appraisal of the structural funds. *Regional Studies*, 29(8), 745–751. <https://doi.org/10.1080/00343409512331349353>
- Bachtler, J., & Wren, C. (2006). Evaluation of European Union Cohesion policy: Research questions and policy challenges. *Regional Studies*, 40(2), 143–153. <https://doi.org/10.1080/00343400600600454>
- Bachtrögl, J., Fratesi, U., & Perucca, G. (2020). The influence of the local context on the implementation and impact of EU Cohesion Policy. *Regional Studies*. Taylor & Francis, 54(1), 21–34. <https://doi.org/10.1080/00343404.2018.1551615>
- Barca, F. (2009). *An agenda for a reformed cohesion policy*. Bruxelles. Available at: https://ec.europa.eu/regional_policy/archive/policy/future/barca_en.htm
- Barca, F., & McCann, P. (2011). *Outcome Indicators and Targets - towards a new system of monitoring and evaluation in EU cohesion policy*. Available at: https://ec.europa.eu/regional_policy/sources/docgener/evaluation/doc/performance/outcome_indicators_en.pdf
- Barlas, Y. (1996). Formal aspects of model validity and validation in system dynamics. *System Dynamics Review*, 12(3), 183–210. [https://doi.org/10.1002/\(SICI\)1099-1727\(199623\)12:3<183::AID-SDR103>3.0.CO;2-4](https://doi.org/10.1002/(SICI)1099-1727(199623)12:3<183::AID-SDR103>3.0.CO;2-4)
- Batterbury, S. C. E. (2006). Principles and purposes of European Union Cohesion policy evaluation. *Regional Studies*, 40(2), 179–188. <https://doi.org/10.1080/00343400600600504>
- Black, L. (2013). When visuals are boundary objects in system dynamics work. *System Dynamics Review*, 29(2), 70–86. <https://doi.org/10.1002/sdr>
- Boselli, S., & Tagle, L. (2003). News from the community. *Evaluation*, 9(2), 229–233. <https://doi.org/10.1177/1356389003009002008>
- Brunazzo, M. (2016). The history and evolution of cohesion policy. In S. Piattoni, & L. Polverari (Eds.), *Handbook on cohesion policy in the EU* (pp. 17–35). Edward Elgar Publishing. <https://doi.org/10.4337/9781784715670.00014>
- Cartwright, A., & Batory, A. (2012). Monitoring committees in cohesion policy: Overseeing the distribution of structural funds in Hungary and Slovakia. *Journal of European Integration*, 34(4), 323–340. <https://doi.org/10.1080/07036337.2011.595486>
- Corte dei Conti. (2017). *Relazione annuale 2017: I rapporti finanziari con l'Unione Europea e l'utilizzazione dei Fondi comunitari*. Rome: Presidenza del Consiglio dei Ministri.
- Cunico, G., Aivazidou, E., & Mollona, E. (2020). European Cohesion Policy performance and citizens' awareness: A holistic system dynamics framework. *Investigaciones Regionales*, 46, 131–162. <https://doi.org/10.38191/iirr-jorr.20.007>
- Davies, S., & Polverari, L. (2011). Financial accountability and European Union cohesion policy. *Regional Studies*. Routledge, 45(5), 695–706. <https://doi.org/10.1080/00343404.2010.529118>
- Doyle, J. K., & Ford, D. N. (1999). Mental models concepts revisited: Some clarifications and a reply to Lane. *System Dynamics Review*, 15(4), 411–415. [https://doi.org/10.1002/\(SICI\)1099-1727\(199924\)15:4<411::AID-SDR181>3.0.CO;2-R](https://doi.org/10.1002/(SICI)1099-1727(199924)15:4<411::AID-SDR181>3.0.CO;2-R)
- Earl, S., Carden, F., & Smutylo, T. (2001). *Outcome mapping*. Ottawa, Canada: International Development Research Centre.
- European Commission (2003-2018). Structural Funds: Report on the Evolution of Payments – Annual series for ERDF and ESF during FP 2000-2006, FP 2007-2013, and FP 2014-2020. Brussels European Commission.
- European Commission. (2013). *Cohesion policy: Strategic report 2013 on programme implementation 2007-2013*. Available at: https://ec.europa.eu/regional_policy/en/information/publications/reports/2013/cohesion-policy-strategic-report-2013-on-programme-implementation-2007-2013
- European Commission. (2014). *An introduction to EU cohesion policy 2014-2020*. Available at: Bruxelles: European Commission https://ec.europa.eu/regional_policy/sources/docgener/informat/basic/basic_2014_en.pdf
- European Commission. (2018a). *Final proposal for a regulation of the European parliament and of the council laying down common provisions on the European regional development fund, the European social fund plus, the cohesion fund, and the European maritime and fisheries fund and finan*. Available at: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2018%3A375%3AFIN>
- European Commission. (2018b). *Programming period 2014-2020: Monitoring and evaluation of European cohesion policy*. Available at: Bruxelles: European Commission <https://ec.europa.eu/sfc/en/system/files/ged/ESFmonitoringandevaluationguidance.pdf>
- European Commission. (2019). *Simplification Handbook. 80 simplification measures in Cohesion Policy 2021-2027*. Available at: https://ec.europa.eu/regional_policy/sources/docgener/factsheet/new_cp/simplification_handbook_en.pdf
- European Court of Auditors. (2016). *Gold-plating*.
- European Court of Auditors. (2018). *Commission's and Member States' actions in the last years of the 2007-2013 programmes tackled low absorption but had insufficient focus on results*. Luxembourg: European Court of Auditors. doi: 10.2865/1131.
- European Parliament. (2011). *Report on absorption of structural and cohesion funds: Lessons learnt for the future cohesion policy of the EU (2010/2305 (INI))*. Bruxelles: European Parliament.

- European Parliament and the Council of European Union. (2013). *Regulation (EU) N° 1303/2013*. Available at: Official Journal of the European Union. Bruxelles: European Parliament and the Council <http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32013R1303>.
- Forrester, J. (1961). *Industrial dynamics*. Massachusetts Institute of Technology Press.
- Forrester, J., & Senge, P. (1980). Tests for building confidence in system dynamics models. *TIMS Studies in the Management Sciences*, 14(1), 209–228.
- Fratesi, U., & Wislade, F. G. (2017). The impact of European Cohesion Policy in different contexts. *Regional Studies*. Taylor & Francis, 51(6), 817–821. <https://doi.org/10.1080/00343404.2017.1326673>
- Gaffey, V. (2013). A fresh look at the intervention logic of Structural Funds. *Evaluation*, 19(2), 195–203. <https://doi.org/10.1177/1356389013485196>
- Gandolfo, D. (2014). *I Fondi Strutturali Europei 2007-2013 in Italia: Un'Analisi delle criticità*. University of Bologna.
- Gates, E. F. (2016). Making sense of the emerging conversation in evaluation about systems thinking and complexity science. *Evaluation and Program Planning*. Elsevier Ltd, 59, 62–73. <https://doi.org/10.1016/j.evalprogplan.2016.08.004>
- Gates, E. F. (2017). Learning from seasoned evaluators: Implications of systems approaches for evaluation practice. *Evaluation*, 23(2), 152–171. <https://doi.org/10.1177/1356389017697613>
- Gore, T., & Wells, P. (2009). Governance and evaluation: The case of EU regional policy horizontal priorities. *Evaluation and Program Planning*, 32(2), 158–167. <https://doi.org/10.1016/j.evalprogplan.2008.10.007>
- Groesser, S. N., & Schwanger, M. (2012). Contributions to model validation: Hierarchy, process, and cessation. *System Dynamics Review*, 28(2), 157–181. <https://doi.org/10.1002/sdr.1466>
- Hassmiller Lich, K., et al. (2017). Extending systems thinking in planning and evaluation using group concept mapping and system dynamics to tackle complex problems. *Evaluation and Program Planning*. Elsevier Ltd, 60, 254–264. <https://doi.org/10.1016/j.evalprogplan.2016.10.008>
- Hooghe, L. (1996). *Cohesion policy and European integration: Building multi-level governance*. Oxford University Press.
- Hummelbrunner, R. (2011). Systems thinking and evaluation. *Evaluation*, 17(4), 395–403. <https://doi.org/10.1177/1356389011421935>
- Kváča, V., & Kokeš, R. (2018). *Systems thinking for European structural and investment funds management*. Bruxelles. <https://doi.org/10.2767/659623>
- Lee, H., et al. (2011). R&D performance monitoring, evaluation, and management systems: A model and methods. *International Journal of Innovation and Technology Management*, 08(02), 295–313. <https://doi.org/10.1142/S0219877011002301>
- Lion, C., & Martini, P. (2006). The evaluation of a Complex Social Program: Lessons learned from the experience of the European Social Fund. *Evaluation and Program Planning*, 29(1), 1–9. <https://doi.org/10.1016/j.evalprogplan.2005.06.001>
- Lion, C., Martini, P., & Volpi, S. (2004). The evaluation of European social fund programmes in a new framework of multilevel governance: The Italian experience. *Regional Studies*, 38(2), 207–212. <https://doi.org/10.1080/0034340042000190172>
- Marchante, A. J., & Ortega, B. (2010). Evaluating efficiency in the implementation of structural funds operations. *Evaluation*, 16(2), 193–209. <https://doi.org/10.1177/1356389009360477>
- Marra, M. (2018). The ambiguities of performance-based governance reforms in Italy: Reviving the fortunes of evaluation and performance measurement. *Evaluation and Program Planning*. Elsevier Ltd, 69, 173–182. <https://doi.org/10.1016/j.evalprogplan.2017.02.006>
- Martinez-Moyano, I. J. (2012). Documentation for model transparency. *System Dynamics Review*, 28(2), 199–208. <https://doi.org/10.1002/sdr.1471>
- Masana, R. E., & Fernández, T. (2019). Monitoring S3: Key dimensions and implications. *Evaluation and Program Planning*. Elsevier, 77(July), 101720. <https://doi.org/10.1016/j.evalprogplan.2019.101720>
- Mceldowney, J. J. (1991). Evaluation and European regional policy. *Regional Studies*. CARFAX PUBL CO PO BOX 25, ABINGDON, OXFORDSHIRE, ENGLAND OX14 3UE, 25 (3), 261–265.
- Meadows, D. (1998). *Indicators and information systems for sustainable development, a report to the balaton group*. Hartland.
- Midgley, G. (2008). Response to paper “Systems thinking” by D. Cabrera et al.: The unification of systems thinking: Is there gold at the end of the rainbow? *Evaluation and Program Planning*, 31(3), 317–321. <https://doi.org/10.1016/j.evalprogplan.2008.04.002>
- Moore, G. F., et al. (2019). From complex social interventions to interventions in complex social systems: Future directions and unresolved questions for intervention development and evaluation. *Evaluation*, 25(1), 23–45. <https://doi.org/10.1177/1356389018803219>
- Naldini, A. (2018). Improvements and risks of the proposed evaluation of Cohesion Policy in the 2021–27 period: A personal reflection to open a debate. *Evaluation*, 24 (4), 496–504. <https://doi.org/10.1177/1356389018804261>
- Nigohosyan, D., & Vutsova, A. (2018). The 2014–2020 European Regional Development Fund Indicators: The Incomplete Evolution. *Social Indicators Research*. Springer Netherlands, 137(2), 559–577. <https://doi.org/10.1007/s11205-017-1610-8>
- OECD. (2015). *Open government data review of Poland*. <https://doi.org/10.1787/9789264241787-en>
- Olaya, C. (2015). Cows, agency, and the significance of operational thinking. *System Dynamics Review*, 31(4), 183–219. <https://doi.org/10.1002/sdr.1547>
- Olejniczak, K., Raimondo, E., & Kupiec, T. (2016). Evaluation units as knowledge brokers: Testing and calibrating an innovative framework. *Evaluation*, 22(2), 168–189. <https://doi.org/10.1177/1356389016638752>
- Perrin, B. (2011). What is a results/performance-based delivery system? An invited presentation to the European Parliament. *Evaluation*, 17(4), 417–424. <https://doi.org/10.1177/1356389011420211>
- Piattoni, S., & Polverari, L. (2016). *Handbook on cohesion policy in the EU*. Edward Elgar Publishing. <https://doi.org/10.4337/9781784715670>
- Polverari, L. (2015). Does devolution increase accountability? Empirical evidence from the implementation of European Union cohesion policy. *Regional Studies*. Taylor & Francis, 49(6), 1074–1086. <https://doi.org/10.1080/00343404.2014.1001351>
- Rainoldi, A. (2010). ‘Dal 1989 al 2009. Vent’anni di intervento del “Fondo europeo di sviluppo regionale” nel Mezzogiorno d’Italia: scelte allocative, utilizzo delle risorse e ruolo del fattore-tempo’. *Rivista giuridica del Mezzogiorno, Trimestrale della Svinez*, 1, 7–42. <https://doi.org/10.1444/32201>
- Rogers, P. J. (2008). Using programme theory to evaluate complicated and complex aspects of interventions. *Evaluation*, 14(1), 29–48. <https://doi.org/10.1177/1356389007084674>
- Saunders, M. (2011). Capturing effects of interventions, policies and programmes in the European context: A social practice perspective. *Evaluation*, 17(1), 89–102. <https://doi.org/10.1177/1356389010394480>
- Smeriglio, A., et al. (2015). *Administrative capacity-building and EU cohesion policy I: Literature review*.
- Smutylo, T. (2001). Crouching impact, hidden attribution: Overcoming threats to learning in development programs. *Evaluation unit*. Ottawa, Canada. Available at: http://www.outcomemapping.ca/download.php?file=/resource/files/admin_en_Crouching_Impact_Hidden_Attribution.pdf.
- Stephenson, P. (2013). Twenty years of multi-level governance: “Where does it come from? What is it? where is it going?” *Journal of European Public Policy*, 20(6), 817–837. <https://doi.org/10.1080/13501763.2013.781818>
- Stephenson, P. (2016). The institutions and procedures of cohesion policy. In S. Piattoni, & L. Polverari (Eds.), *Handbook on cohesion policy in the EU* (pp. 36–49). Edward Elgar Publishing. <https://doi.org/10.4337/9781784715670.00015>.
- Sterman, J. (2000). *Business dynamics: Systems thinking and modeling for a complex world*. Boston: Irwin/McGraw-Hill (McGraw-Hill Higher Education).
- Sterman, J., & Rahmandad, H. (2012). Reporting guidelines for simulation-based research in social sciences. *System Dynamics Review*, 28(4), 396–411. <https://doi.org/10.1002/sdr.1481>
- Unicef. (1991). *A UNICEF guide for monitoring and evaluation making a difference?* UNICEF. Evaluation Office.
- Vennix, J. A. M. (1996). *Group model building*. Available at: Chichester: Wiley <https://books.google.nl/books?id=F4ciAQAAAJ>.
- Viñas, V. (2009). The European Union’s drive towards public policy evaluation. *Evaluation*, 15(4), 459–472. <https://doi.org/10.1177/1356389009341900>
- Wholey, J. S. (1983). *Evaluation and effective public management*. Little, Brown Boston.
- Wholey, J. S. (1996). Formative and summative evaluation: Related issues in performance measurement. *The American Journal of Evaluation*, 17(2), 145–149. <https://doi.org/10.1177/109821409601700206>
- Wojtowicz, D. M., & Kupiec, T. (2018). Reluctant to learn? The use of evaluation to improve EU cohesion policy implementation in Polish and Spanish regions. *Evidence and Policy*, 14(1), 103–120. <https://doi.org/10.1332/174426416X14788831396548>

Mr Giovanni Cunico has worked as a Research Fellow in the project PERCEIVE, an EU-funded action aiming at improving policy-makers’ understanding of the EU regional policy implementation and communication, at the University of Bologna (Italy). After obtaining a Bachelor in Political Science at the University of Trento (Italy), he graduated from the European Master in System Dynamics. He is interested in applying system dynamics and participatory approaches to policy development environments, grand challenges, and sustainability-related issues. ORCID: <https://orcid.org/0000-0003-0450-7725>

Dr Eirini Aivazidou served as a Research Fellow in System Dynamics at the Department of Computer Science and Engineering of the University of Bologna, Italy, in the Horizon 2020 project PERCEIVE. She is currently an Adjunct Lecturer at the School of Mechanical Engineering of the Aristotle University of Thessaloniki (AUTH), Greece. She holds a Ph.D. in Supply Chain Water Footprint Management (funded by the Onassis Foundation), an M.Sc. in Transportation Engineering, and a Diploma in Mechanical Engineering, all with Distinction from AUTH. Her research work has been published in several peer-reviewed journals and scientific books. ORCID: <https://orcid.org/0000-0002-4208-2361>

Dr Edoardo Mollona graduated with cum laude in Strategic Management from the Bocconi University in Milan (Italy) and received a Ph.D. in Strategic Management/Decision Sciences from the London Business School (UK). He is a Full Professor in Business Economics and Policy Modelling at the Department of Computer Science and Engineering at the University of Bologna (Italy). He has published books and journal articles, he has been in chairs of international conferences, he edited the book “Computational Analysis of Firms’ Strategy and Organizations” (Routledge), and he served as the principal investigator of the EU-funded Horizon 2020 project PERCEIVE. ORCID: <http://orcid.org/0000-0001-9496-8618>