EUROPEAN
COMMISSION

Brussels, 10.7.2023
SWD(2023) 240 final

COMMISSION STAFF WORKING DOCUMENT

IMPACT ASSESSMENT REPORT

Accompanying the document

Proposal for a Regulation of the European Parliament and of the Council

amending Regulation (EC) No 223/2009 on European statistics

{COM(2023) 402 final} - {SEC(2023) 269 final} - {SWD(2023) 241 final}

1. INTRODUCTION: POLITICAL AND LEGAL CONTEXT 1
1.1. European statistics 1
1.2. Changing societal, political and legal context 2
2. PROBLEM DEFINITION 5
2.1. What is/are the problems? 5
2.2. What are the problem drivers? 7
2.3. How likely is the problem to persist? 10
3. WHY SHOULD THE EU ACT? 11
3.1. Legal basis 11
3.2. Subsidiarity: Necessity of EU action 11
3.3. Subsidiarity: Added value of EU action 12
4. OBJECTIVES: WHAT IS TO BE ACHIEVED? 13
4.1. General objective 13
4.2. Specific objectives 13
5. WHAT ARE THE AVAILABLE POLICY OPTIONS? 15
5.1. What is the baseline from which options are assessed? 15
5.2. Description of the policy options 16
5.3. Options discarded at an early stage 23
6. WHAT ARE THE IMPACTS OF THE POLICY OPTIONS? 23
6.1. Introduction 23
6.2. Impact of policy option 0: the baseline option 25
6.3. Impact of policy option 1: the first legislative option 27
6.4. Impact of policy option 2: the second legislative option 31
7. HOW DO THE OPTIONS COMPARE? 34
7.1. Effectiveness 34
7.2. Efficiency 35
7.3. Coherence 36
7.4. Feasibility 37
8. PREFERRED OPTION 38
8.1. The preferred option and its implementation 38
8.2. Estimated impact of the preferred option 41
8.3. REFIT (simplification and improved efficiency) 42
8.4. Application of the ‘one in, one out’ approach 43
8.4.1. Potential new burden on citizens 43
8.4.2. Potential new costs on businesses 43
9. HOW WILL ACTUAL IMPACTS BE MONITORED AND EVALUATED? 44
ANNEX 1: PROCEDURAL INFORMATION 48
1. LEAD DG, DECIDE PLANNING/CWP REFERENCES 48

2. ORGANISATION AND TIMING 48
3. CONSULTATION OF THE RSB 48
4. EVIDENCE, SOURCES AND QUALITY 53
ANNEX 2: STAKEHOLDER CONSULTATION (SYNOPSIS REPORT) 56
1. Introduction 56
2. Key stakeholders 56
3. Summary of results 58
ANNEX 3: WHO IS AFFECTED AND HOW? 66
1. PRACTICAL IMPLICATIONS OF THE INITIATIVE 66
2. SUMMARY OF COSTS AND BENEFITS 69
3. RELEVANT SUSTAINABLE DEVELOPMENT GOALS 82
ANNEX 4: ANALYTICAL METHODS 83
ANNEX 5: LEGAL CONTEXT 86
ANNEX 6: ASSESSED USE CASES 90
ANNEX 7: DATA GAPS IN TERMS OF CONTENT, FREQUENCY, TIMILINESS AND GRANULARITY OF EUROPEAN STATISTICS 95
ANNEX 8: SME TEST 100

Table 1: Policy options and measures 18
Table 2: Comparison of quantitative impact of policy options 36
Table 3: Comparison of policy options 38
Table 4: Expected REFIT Cost Savings under the preferred option PO1 42
Table 5: Monitoring of impacts 44
Table 6: Stakeholder participation 57
Table 7: Results of targeted consultations 62
Table 8: Stakeholders' estimate of costs in various use cases 63
Table 9: Cost and benefit for PO0: Reuse of new data sources and crisis response 73
Table 10: PO0: Cost and benefit: Voluntary ESS data sharing 74
Table 11: PO1: Estimated differential cost and benefit for data reuse and crisis response 75
Table 12: PO1: Estimated differential benefits and costs of the measure on mandatory ESS data sharing 76
Table 13: Overview of benefits – preferred option 77
Table 14: Overview of costs – preferred option 78
Table 15: PO1 – Distribution of costs and benefits between statistical offices and data holders 78
Table 16: PO2 – estimated differential costs and benefits for data reuse and crises response 80

Table 17: PO2 – Distribution of costs and benefits between statistical offices and data holders 80
Table 18: PO2 - Estimated differential benefits of the measure on mandatory ESS data sharing 81
Table 19: PO2 - Estimated costs of additionally providing access to shared raw data for research purposes 81
Table 20: Overview of relevant Sustainable Development Goals – Preferred Option..... 82

Figure 1: The problem definition 5
Figure 2: The intervention logic 13
Figure 3: PO0: Application of statistical products / use cases 73
Figure 4: PO1: Application of statistical products / use cases 75
Figure 5: PO2 - Application of statistical products / use cases 79
Figure 6: Visual representation of B2G data sharing chains for official statistical production, and of the benefits and the B2B data sharing chains they generate 85
Figure 7: Model of stakeholders related to the ESS [source: ICF]..... 89

Glossary

Term or acronym	Meaning or definition
AWP	Annual Work Programme for European statistics
B2B	Business to business data sharing
B2G	Business to government data sharing for the public interest
B2G4S	Business to government data sharing for the purpose of generating official statistics
CEDS	Common European Data Spaces
DGA	Data Governance Act
ECA	European Court of Auditors
EDIB	European Data Innovation Board
ESP	European Statistical Programme
European Statistical System (ESS)	The partnership between the Community statistical authority, which is the Commission (Eurostat), and the NSIs and other national authorities responsible in each Member State for the development, production and dissemination of European statistics
GDPR	General Data Protection Regulation
Internet of Things (IoT)	A network of physical devices, vehicles, home appliances and other items embedded with connectivity software, which enables these objects to connect and exchange data
LFS	EU Labour Force Survey
MNO	Mobile Network Operator
National Statistical Institute (NSI)	The national statistical authority designated by a Member State as the body having the responsibility for coordinating all activities at national level for the development, production and dissemination of European statistics
PET	Privacy-enhancing techniques
REFIT	Regulatory Fitness and Performance Programme of the European Commission
SCM	Standard Cost Model
SILC	EU Statistics on Income and Living Conditions
SMEI	Single Market Emergency Initiative

1. INTRODUCTION: POLITICAL AND LEGAL CONTEXT

1.1. European statistics

This impact assessment accompanies the proposal for a Regulation amending Regulation (EC) No 223/2009 of the European Parliament and of the Council of 11 March 2009 on European statistics. This regulation provides the legal framework at EU level for the development, production and dissemination of European statistics.

Since 1953 along with the evolution of the European Union, European statistics have played an increasingly important role for underpinning the Union activities, policies and legislative acts from their design and implementation to their monitoring and evaluation. Today, they are an indispensable element of the European democratic society and a fundamental tool for fostering transparency and accountability on policymaking at all levels of public administration. European statistics are, therefore, by their very nature, an invaluable and consensual public good[1], provided to everybody at the same time and for free. They are reliable and impartial and open to public scrutiny, and they must comply adequately with internationally agreed quality and other professional standards, such as privacy and confidentiality.

European statistics are provided by the **European Statistical System** (ESS) that is a partnership between Eurostat and the national statistical institutes (NSIs), as well as other national authorities responsible for the development, production and dissemination of European statistics in each Member State. **Eurostat** is the statistical authority of the Union that leads the development of European statistics in response to evolving policy needs at EU level; coordinates statistical production based on common EU statistical standards, methods, procedures, practices, and tools; and assesses the quality of data transmitted by the Member States. It also coordinates on statistical matters with other European Institutions and bodies and at global level with international organisations such as the United Nations (UN), the International Monetary Fund (IMF) and the Organisation for Economic Co-operation and Development (OECD).

European statistics are determined in the European Statistical Programme (ESP), the duration of which coincides with the Multiannual Financial Framework. The current ESP is an integral part of the Single Market Programme 2021-27. They are developed, produced and disseminated in conformity with the statistical principles enshrined in Article 338(2) of the Treaty on the Functioning of the European Union and specified in Regulation (EC) No 223/2009. The set of statistical principles is further elaborated in the European statistics Code of Practice that is progressively adapted to the fast evolving institutional, policy and professional environment within which development, production and dissemination of European statistics takes place.

[1] Cf European Court of Auditors' (ECA) press release of 29 November 2022, "European statistics should better meet user needs", where the ECA member, Ildikó Gáll-Pelcz, is quoted for stating that "[Statistics] are a public good, and must be generated first and foremost with users in mind. In an age of disinformation and serial crises, it is paramount that European official statistics must be high-quality, meet users' needs and explore innovative ways of production."

1.2. Changing societal, political and legal context

Digital transformation and changing society

The ESS operates in a fast-evolving world marked by digital transformation. Since the adoption of the Regulation (EC) No 223/2009 the society has changed profoundly, and at an unprecedented speed. The amount of data generated yearly is counted in zettabytes[2]. Data, which are to a large extent created as a by-product of digital services and connected devices, contain a huge potential as an essential infrastructural resource for economic growth, innovation and the overall well-being of society[3]. Data are becoming a critical resource for start-ups and SMEs' growth as well as for achieving the Green Deal objectives.

Together with the 'twin' green transition, the digital transformation has led to the emergence of new business models (e.g., platform economy) and of new digital services, as well as to profound changes in labour markets and well-being. Such transformations are fed with, and are the source of, large amounts of data that have the potential to be used as primary input for compiling official European statistics through adequate methods and modern data technologies. In some cases, data that are the by-product of digital services are the only data sources that can be used to measure emerging economic phenomena brought by digital transformation such as the platform economy.

Growing demands for timelier and more granular European statistics especially in times of crises

Digitalization and the emergence of big data and the Internet of Things have brought a growing availability of real-time and close to real-time data for decision-making, that in turn have opened opportunities for providing real-time views of the economies, labour markets and societies. Additionally, big data provide opportunities for statistics and statistical insights with more granularity to paint the picture of various social groups (e.g. youth, women and migrants) and regional entities and possibly foretell changes for the years to come.

These trends have raised questions about the relevance of the economic and social data and indicators by official statistics, provided usually on a quarterly and monthly basis and with a certain lag after the end of the reference period. High-frequency and close to real-time data appeared to be critical for example when it came to track the rapid economic disruption brought by the Covid-19 pandemic. Furthermore, European statistics on labour markets reported on a quarterly and monthly basis had struggled to keep pace with the wave of unemployment and growth in the labour market slack, accompanied by government furlough and other schemes.

Overall, the recent financial, migration and Covid-19 crises, followed by the Russian military aggression against Ukraine, have amplified these demands and expectations for timelier and more detailed European statistics, which are needed to ensure the best EU response to crises and support informed decisions.

---

[2] 1 zettabyte (ZB) = 10^{21}.

[3] This is elaborated in, e.g., the report of the High-Level Expert Group on Business-to-Government Data Sharing: *Towards a European strategy on business-to-government data sharing for the public interest.*

The growing demands for timelier, more frequent, and more granular data, has made it necessary for Eurostat and the national statistical authorities to establish data partnerships and engage continuously in collaboration with partners from both public and private sector for the reuse of new and innovative data sources by removing barriers of reuse of these sources, while ensuring full respect of privacy and confidentiality.

Over the recent past years, the ECOFIN Council has consistently addressed in its annual Conclusions on European statistics the need for the ESS to explore new data sources and technologies and develop innovative methods for producing high-quality, richer and timelier European statistics. Issues related to a sustainable access to new data sources, enhanced data sharing and data linking, improved efficiency and burden reduction, and increased agility of the system, especially in times of crises, have been particularly highlighted. Most recently, in the **ECOFIN Council conclusions on EU statistics of 8 November 2022**[4], the Council stated that it was looking forward to a possible proposal by the Commission on the revision of Regulation (EC) No 223/2009 on European statistics, addressing in particular the issue of access to new data sources from the specific perspective of European statistics, as well as ways to ensure increased agility and responsiveness of the ESS to policy needs.

These issues have been also systematically raised in the context of the annual statistical dialogue of Director General of Eurostat with the **European Parliament Committee on Economic and Monetary Affairs**. Furthermore, they have been consistently raised by the **European Statistical Governance Advisory Board** (ESGAB) and the **European Statistics Advisory Committee** (ESAC), which are two independent statistical bodies at EU level, as well as in the recent European Court of Auditors special report on the relevance of European Statistics. In particular, the European Statistical Governance Advisory Board highlighted in its 2021 Annual Report[5] the potential of new digital data sources to contribute to the objectives of relevance, accuracy, and timeliness, and to reduce the burden on respondents and increase cost-effectiveness. It also identified the need for Eurostat to be able to react quickly to unexpected and urgent statistical demands for policymaking. In the same vein, the **European Court of Auditors** noted, when recently releasing its Special Report 26/2022[6], that the ESS was not flexible enough to respond quickly with new sets of data when new needs arise.

Finally, as part of the overall policy context, it is important to note that the Presidents and Directors-General of the NSIs have also consistently called for the ESS to be equipped with the necessary regulatory instruments and innovative tools to enhance the access, use and integration of all data available and allow the production of multi-source European statistics. For instance, as an outcome of the high-level meeting organised in Lyon in April 2022 under the French Presidency and of a subsequent dedicated meeting in Luxembourg in May 2022, they reached a large consensus on a number of topics considered valuable in the revision of the Regulation, not only on the need for sustainable access to new data sources, but also on the recognition of the possibility for NSIs to assume new roles and tasks in the emerging data ecosystems, and the fostering of data sharing in the ESS.

---

[4] Council adopted conclusions on statistics - Consilium (europa.eu)
[5] ESGAB Annual report 2021 (europa.eu)
[6] Special report 26/2022: European statistics – Potential to further improve quality (europa.eu)

Recent EU policy and legal initiatives

The overall political context is also marked by the recent legislative initiatives that have been undertaken as part of the **European Strategy** for **Data** put forward by the Commission in February 2020[7] with the aim to strengthen Europe's position globally by making better use of data driven innovation. These initiatives – **Data Governance Act (DGA)**[8] and the **proposal for a Data Act**[9] establish provisions that increase data availability and data reuse, establish data governance and open room for new actors in the emerging European and national data ecosystems. These are developing rapidly and constitute an opportunity for European statistics producers to harness the power of digital data for high-quality statistics in response to unmet users' demands. The evolution of these data ecosystems imply that large amounts of data can be available as input for producing European statistics using adequate methods and modern data technologies.

The **INSPIRE Directive**[10], that has established an infrastructure for spatial information in Europe to support environmental policies and the **Interoperable Europe Act** proposal[11] that is aiming to strengthen cross-border interoperability and cooperation in the public sector across the EU, are also part of the legal context to consider for this initiative as well as the current proposal for a **Single Market Emergency Initiative (SMEI)**[12] that foresees to allow the Commission to receive targeted information for official statistics, but only from economic operators in the crisis-relevant supply chains.

Finally, special mention should be made to the **General Data Protection Regulatio**n **(GDPR)**[13] and the **ePrivacy Directive**[14], which provide a solid and trusted legal framework for the protection of personal data at EU level.

A more detailed overview of the existing and very dynamic legal context is presented in Annex 5.

---

[7] See https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_en

[8] Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724.

[9] Proposal for a Regulation of the European Parliament and of the Council on harmonised rules on fair access to and use of data (Data Act).

[10] Cf. Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE), OJ L *108, 25.4.2007, p. 1*.

[11] Commission proposal for a Regulation of the European Parliament and of the Council laying down measures for a high level of public sector interoperability across the Union (Interoperable Europe Act), COM(2022) 720 final, adopted on 18 November 2022.

[12] See Proposal for a Regulation of the European Parliament and of the Council establishing a Single Market emergency instrument and repealing Council Regulation No (EC) 2679/98.

[13] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)

[14] Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications This directive is expected to be eventually repealed by the Regulation on Privacy and Electronic Communications based on a proposal by the Commission (COM/2017/010) currently under discussions between the European Parliament and the Council.

2. PROBLEM DEFINITION[15]

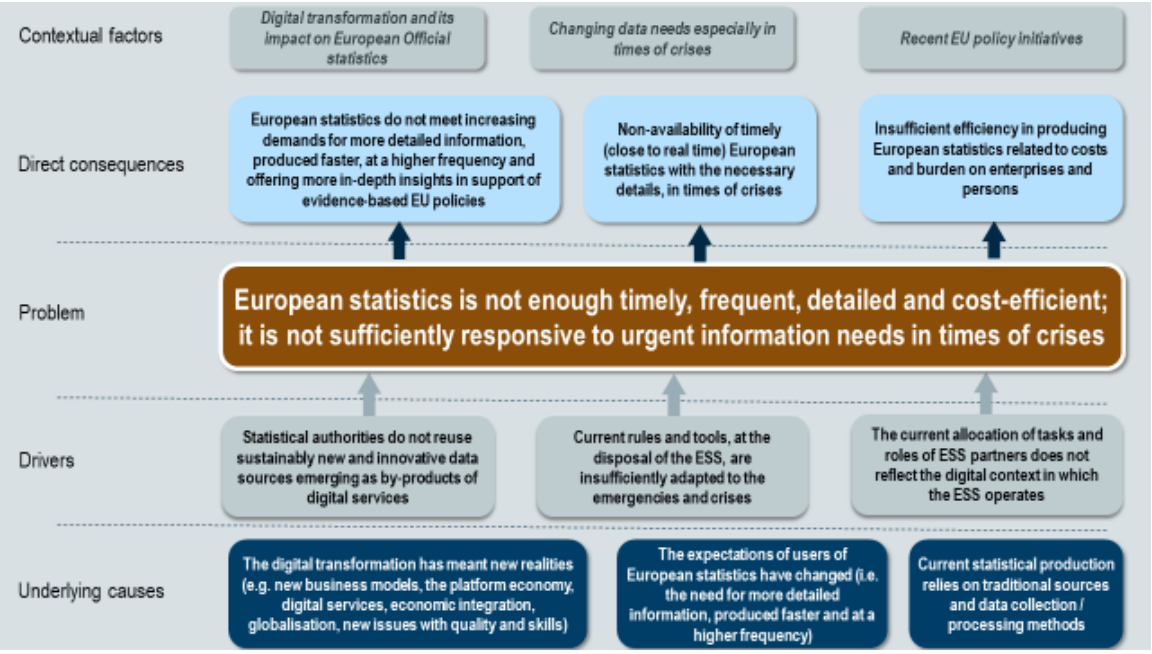

Figure 1: The problem definition

2.1. What is/are the problems?

The problem that this initiative is aiming to address is that European statistics is not enough timely, frequent, detailed and cost-efficient; it is not sufficiently responsive to urgent information needs in times of crises**.**

European statistics are developed, produced and disseminated according to the principles, mechanisms, tools and governance laid down in Regulation (EC) No 223/2009. The latter was developed in the early 2000s. Therefore, it reflects the way statistics were produced at that time, almost fully based on sample surveys, population and other censuses and administrative records held by public authorities. The legal framework and the statistical practice based on it do not reflect the new realities brought by digital transformation such as new data sources and technologies. As a consequence, the ESS is facing growing difficulties to fill the emerging data gaps and meet user demands in terms of timeliness and frequency of providing those statistics as well as of granularity at which they are needed to make informed policy decisions (e.g. decomposed by social groups and regional and local dimensions).

Recent crises, in particular Covid-19, have shown that while the ESS has demonstrated strong resilience and continued producing and disseminating traditional statistics according to the established deadlines, it has not responded quickly enough and has not

[15] The problem definition has been elaborated in more detail by ICF, in their Study to support an impact assessment for the revision of Regulation (EC) No 223/2009 on European statistics, final report. Later references to this study are denoted by [ICF]. That study also includes an assessment of stakeholders of the problem. See also Annex 1.4 for evidence used and Annex 2 on stakeholders consultation.

met the new urgent information needs. It is essential that the ESS is agile and effective in its response to urgent information needs arising in times of crises and following emergency mechanisms established by EU law such as those on Public Health[16] or Migration[17].

Going beyond the traditional data sources used for compiling European statistics and especially through reusing of data already collected for other purposes has a potential to reduce costs and administrative burden on businesses and people. This potential is currently not sufficiently utilised, thus leaving room for improving efficiency of producing European statistics.

The problem entails the following direct consequences:

- European statistics do not meet increasing demands for more detailed information, produced faster, at a higher frequency and offering more in-depth insights in support of evidence-based EU policies.
- There is non-availability of timely (close to real time) European statistics, with the necessary details, in times of crises.
- There is insufficient efficiency in producing European statistics regarding the costs and burden on enterprises and persons.

Over recent years, a solid evidence has been accumulated about the existence of the problem of insufficient timeliness, frequency, granularity and cost-efficiency of European statistics as well as of its inadequate responsiveness to urgent information needs in times of crises**.**

European statistics are determined by the ESP, and each ESP is evaluated as stipulated in Regulation (EC) No 223/2009. The most recent evaluation of European pertains to the ESP for the period 2013 – 2020[18]. The evaluation report covers the entire period of the programme 2013-2020, and it is accompanied by a European Commission staff working document. The evaluation study was conducted by an external contractor to give an independent opinion. To get robust results, the contractor used different sources, starting with the review of existing documents. Then, the contractor carried out an extensive consultation with users and producers of statistics. This included notably i) scoping interviews with representatives of Eurostat and other Directorates-General of the Commission; ii) a public consultation; iii) targeted surveys of users and producers; and iv) 50 interviews with different types of stakeholders. The contractor also carried out four thematic case studies and five country case studies.

The evaluation report concluded that "while the ESP implemented appropriate activities to meet its objectives, the analysis showed that these activities were not enough to deliver all the statistics that users had wished for." Moreover, the report pointed to "remaining weaknesses with the timeliness and the completeness of European statistics" and in the remarks related to coherence, it highlighted that: "(…) the lack of flexibility of European statistics to respond to emerging needs (…) might cause a misalignment with other EU

[16] Cf. https://health.ec.europa.eu/system/files/2021-09/hera_2021_decision_en_0.pdf
[17] Cf. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32001L0055
[18] Report from the Commission to the European Parliament and the Council on the final evaluation of the implementation of the European statistical programme 2013-2020, COM(2021) 794 final.

strategies should these needs not be covered." A prominent recommendation is to focus on innovation, new methods and better use and integration of new data sources in order to satisfy the increasing demands for new and timelier statistics, while reducing costs and administrative burden on businesses and citizens in the production of European statistics.

The fact that European statistics are not fit to the digital age and times of crises has also been recognised by a number of bodies concerned with European statistics, as already mentioned in the previous chapter: the ECOFIN Council, the ESGAB, the ESAC, and the European Court of Auditors. In a number of specific areas of statistics, impact assessments have been carried out that corroborate the existence of the problem[19]. During the on-line public consultation, to the question, what needs to be done to make European statistics fit for the future and more relevant to user needs 70% of respondents considered it most important to combine sources to provide more and better insights into economic and societal developments, 66 % to provide more granular statistics (e.g. for social groups and territorial units), and equally 66% to provide more up-to-date statistics, e.g. through flash estimates and more frequent statistics. Only 11% of respondents consider that European statistics are sufficiently responsive to emerging user demand, including during public emergencies and crises, whereas 72% consider that European statistics are somewhat responsive, but not enough and 8% consider them not responsive at all.

Insufficient timeliness and granularity of European statistics have appeared prominently in the last two user satisfaction surveys conducted by Eurostat in 2019 and 2022 respectively. This report substantiates and specifies the problem of European statistics in Annex 7 where an illustration is provided of issues related to important statistical domains, together with their effects, thereby demonstrating the importance of the problem and the expected consequences and impacts of addressing it. Addressing these issues would be beneficial for different stakeholders. Policy makers will be supported in designing implementing and monitoring policies, businesses will have access to detailed information about potential markets, individuals will be better informed and can better participate in the democratic process, research can contribute to society by better informing on problems, and media can use statistics to provide timely information based on trustworthy sources.

2.2. What are the problem drivers?

Three problem drivers have been identified:

- Statistical authorities do not sustainably reuse new and innovative data sources emerging as by-products of digital services.
- Current rules and tools, at the disposal of the ESS, are insufficiently adapted to crises.
- The current allocation of tasks and roles of ESS partners does not reflect the digital context in which the ESS operates.

[19] Examples are the impact assessment for the Regulation on European statistics on population and housing, SWD(2023) 14 final, the impact assessment for the Regulation on European business statistics, SWD(2017) 98 final, and the impact assessment for the Regulation on a common framework for European statistics relating to persons and households, SWD(2016) 283 final.

Problem driver 1: Statistical authorities do not sustainably reuse new and innovative data sources emerging as by-products of digital services

The new and innovative data sources that are by products of digital services or connected devices are mostly privately held. Examples of these sources include mobile phone data, banking services data, smart meters data, on-line job advertisements. Currently, they are used on an ad hoc basis, based on voluntary agreements that are limited in time, i.e., partnerships agreements between data holders and the national statistical authorities. These agreements most often cover one-off projects focused on experimentation and research. New digital data sources are hardly used in the regular statistical production. In many cases, the process for setting up agreements may take at least one year, if not more – for projects limited in time and purposes.

The current legal framework does not cover the reuse of privately held data emerging as by-products of digital services and the Internet of Things for the sake of compiling European statistics. Moreover, the legal environments in Member States vary with respect to business to government data sharing for the compilation of official statistics (B2G4S). Businesses are faced with legal uncertainties The absence of an encompassing legal framework also discourages investing in B2G4S.

The High-level expert group on facilitating the use of new data sources for official statistics concludes that the European statistical system "has struggled to keep up with the new data-rich world, although it has very seriously tried to access new data sources and privately held data in partnership with businesses" (p.16 of the Report[20]). Companies may have little or no economic incentives to share data with public sector organisations and perceive various obstacles:

- Setting up and implementing data-sharing collaborations could be costly and bring additional administrative burden. For example, additional specialized manpower is needed for negotiating a contractual arrangement or setting up the operational elements to make the data reuse possible.
- Companies may be afraid of incurring possible revenue losses and opportunity costs or lose their competitive advantage in upstream or downstream markets because of making their data available to public authorities such as NSIs or Eurostat.
- Data providers may fear that their data can be used to impose new regulations (hence costs and/or additional administrative burdens) on their operations (for example in the case of the regulation of the gig economy).
- Sharing data could involve risks of data leaks or hacks.

The lack of an encompassing legal framework not only means that businesses as well as ESS partners are faced with legal uncertainties, but also that they have to comply with the different rights, obligations and safeguards of the Member States. Most of the current B2G4S initiatives have been established spontaneously and are very diverse and dissimilar, for instance in their 'rules of engagement'. There is also a general lack of

[20] Report of the High-Level Expert Group on facilitating the use of new data sources for official statistics: Empowering society by reusing privately-held data for official statistics – a European approach. References to this report are denoted by [EG B2G4S] in this impact assessment, with B2G4S standing for Business-to-Government for statistics.

transparency about the B2G4S data sharing that takes place, which also contributes to the low awareness of the potential value of such activity.

Problem driver 2: Current rules and tools, at the disposal of the ESS, are insufficiently adapted to crises

Unanticipated events may cause urgent data needs that require quick action at EU level, underpinned by prompt, often close to real-time statistics that are comparable across the Member States. Responding to such needs by the ESS is hampered by the lack of a mechanism to initiate urgent actions at the EU level, since this cannot be addressed within the regular planning framework. The ESS has already strong collaboration mechanisms that take into account the differences among national statistical systems and respect the subsidiarity principle. However, Regulation (EC) No 223/2009 does not provide mechanisms and tools to react fast and in collective coordinated manner in response to urgent information needs in times of crises. The response can include temporary actions such as new data collections or utilising existing data to produce new statistics and provide additional statistical insights. It can also include the creation of temporary task forces or expert groups to develop methodological guidelines or harmonised methodologies for producing urgently statistics that are comparable across the EU member States.

There are several on-going initiatives at EU level to lay down provisions and mechanisms in certain sectors, as well as the proposal for the Single Market Emergency Initiative (SMEI), that aim at establishing a crisis response mechanism. A mechanism that would allow the ESS to react in times of crises will be strictly complementary to and support the other crisis response mechanisms at EU level once these are activated. This means that the ESS will clearly not decide on its own on the existence of a crisis but will take action only in the case a crisis mechanism has been formally triggered by an Institution according to established procedures in Union law.

Problem driver 3: The current allocation of tasks and roles of ESS partners does not reflect the digital context in which the ESS operates.

The digital revolution has brought new opportunities to exchange and share data under a secure infrastructure as well as allowed for linking data across different sources and producing new, deeper insights into society and economy. As regards the ESS, data sharing among its partners is rather limited and fragmented. Multiple obstacles persist in the sharing of data, particularly in transferring micro-level data across borders. This can significantly hinder the collaboration and cooperation of NSIs, and other partners, as demonstrated in the Nordic Mobility project[21]. Furthermore, such differences can be detrimental to fair competition, transparency, and equal treatment of economic operators[22]. With fragmentation across the Member States regarding the modalities and scope of data sharing, the cost of European statistics is negatively affected [EG B2G4S p7]. Issues such as data localization requirements, a lack of standardisation of data

[21] Brun, N., Ekmark, S., Munch Haagensen, K., Harðarson, Ó., Rustad Holseter, A.M., Nome Næsheim, H. and Ruotsalainen, K., 2021. *Nordic Cross-border Statistics: The results of the Nordic Mobility project 2016-2020*. Nordic Council of Ministers.

[22] ESS (2017), Position paper on access to privately-held data which are of public interest.

production and storage all increase the administrative and economic burden on NSIs to both produce and access required data[23].

European statistics on cross-border flows and phenomena will have the highest quality, coverage and detail if the available data can be combined in the compilation of the statistics concerned [EG B2G4S p7]. This requires data sharing, the lack of which also contributes to inefficiencies, as seen in problem driver 2. Examples are migration, labour mobility and foreign investment, where data sharing is essential for getting consistent and correct statistics. The lack of clear, accurate, and available statistical data related to cross-border flows and phenomena inhibits policymakers in making informed decisions[24].

The task distribution among the ESS partners also does not reflect the current digital context. Historically, European statistics have been based on sample surveys, population and other censuses and reuse of administrative records held by public authorities. Data collection took place at national level. Increasing availability of data sources at European or global level raises questions on the efficiency of this model. It might be more efficient to make data from international sources (usually multinationals) available to Eurostat and then allow further reuse by NSIs under strict conditions rather than to have data requests from up to 27 countries, which may not be harmonised. In the latter case, the administrative burden may then also be higher than needed. However, Regulation (EC) No 223/2009 currently does not explicitly provide such a possibility.

Interaction between the problem drivers

The problem drivers are interconnected. The first one, that statistical authorities do not sustainably reuse new data sources emerging as by-products of digital services, interacts with the second one, that current rules and tools, at the disposal of the ESS, are insufficiently adapted to crises. If new data sources can be reused by statistical authorities, this facilitates the response of urgent information needs in times of crises.

The third problem driver, that the current allocation of tasks and roles of ESS partners does not reflect the digital context in which the ESS operates, interacts with the other two problem drivers. Sustainable reuse of new data sources requires an update of the allocation of tasks and roles of ESS partners for situations where such data sources can best be accessed by one ESS partner on behalf of the whole ESS, in particular Eurostat. Similarly, a quick response to crises requires an initiating and coordinating role for Eurostat, which is not recognized in the current allocation of tasks and roles of ESS partners.

2.3. How likely is the problem to persist?

First and foremost, continuing with the current production methods limited to traditional data sources (e.g. surveys and administrative records) and not embracing the opportunities, brought by digital transformation, will make it increasingly difficult and

[23] World Economic Forum, 2020, *A Roadmap for Cross-Border Data Flows: Future-Proofing Readiness and Cooperation in the New Data Economy*.
[24] European Commission, 2017. Boosting growth and cohesion in EU border regions. Communication from the Commission to the Council and the European Parliament.

close to impossible for the ESS to meet user demands for timelier, more frequent and more granular European statistics, even if additional resources are allocated to statistical authorities.

Second, the lack of responsiveness and agility of the ESS to meet emerging user demands especially in times of crises will increase. This will be made even more pronounced by the lack of a crisis response mechanism at the ESS level as well as the general deterioration of the ESS ability to make relevant, detailed and timely statistics in a context marked by fast changing data.

Third, the costs and burden on respondents will not be reduced, rendering production of European statistics more and more inefficient and ineffective in the context of the recent EU policy initiatives related to data.

Fourth, the mismatch between the current and desirable roles of statistical authorities within European and national data ecosystems will increase given that these ecosystems are likely to develop further.

3. WHY SHOULD THE EU ACT?

3.1. Legal basis

Article 338 of the Treaty on the Functioning of the European Union provides the overall legal basis for European statistics. Based on Article 338(1), the European Parliament and the Council, acting in accordance with the ordinary legislative procedure, shall adopt measures for the production of statistics when necessary for the performance of the activities of the Union. Furthermore, Article 338(2) sets out the requirements to produce European statistics, stating that they must conform to standards of impartiality, reliability, objectivity, scientific independence, cost-effectiveness and statistical confidentiality.

Article 338(1) TFEU is thus the legal basis for this initiative.

3.2. Subsidiarity: Necessity of EU action

In order to demonstrate that EU action is necessary to address the problem that European statistics are not enough timely, frequent, detailed and cost-efficient, and that they are not sufficiently responsive to urgent information needs in times of crises, the problem drivers have to be considered.

The first problem driver, i.e., that statistical authorities do not sustainably reuse new data sources emerging as by-products of digital services, is related to the fact that the EU legal framework does not mandate such access and reuse. This is also true for the legal frameworks of the large majority of Member States, as explained in section 2.2, and this is not expected to substantially change in the near future. Moreover, to the extent that reuse of privately held data for official statistics is possible in certain Member States, the conditions and safeguards differ or are ad hoc. As a consequence, reuse of privately held data for statistics tends to be on a voluntary and temporary basis, and thus not sustained.

Thus, addressing the first problem driver necessitates EU action, not only to open the door to reuse of privately held data for official statistics at a European scale, but also to make this reuse sustainable and to harmonise such reuse within the EU. Such

harmonisation would also ensure a level playing field for businesses as data holders. A harmonised approach at EU level would bring legal clarity and ensure a fair treatment of data holders that are active in multiple Member States.

The second problem driver, i.e., that current rules and tools, at the disposal of the ESS, are insufficiently adapted to crises, necessitates EU action, since such situations require coordinated information at the European level. In such situations, action at the European level is needed to initiate and coordinate the production of the information needed.

The third problem driver, i.e., that the current allocation of tasks and roles of ESS partners does not reflect the digital context in which the ESS operates, is cross-border in nature. An optimal allocation would improve efficiency as well as effectiveness, for instance by increased data sharing between ESS partners, or more effective measurement of cross-border phenomena. Since such improvements involve the interaction between ESS partners in different Member States, action at EU level is necessary.

Without action at EU level, the problems that have developed will continue. The existing legislative framework governing European statistics may become less relevant and less effective in achieving its objectives in the years to come as new data sources and technologies emerge. Over time, European statistics may also diverge further from users' needs in terms of content, granularity, desired frequency or timeliness. Finally, Member States' approaches regarding data sharing and use of new data sources will increasingly diverge leading to less comparable statistics, which consequently risks compromising policymaking at EU level.

3.3. Subsidiarity: Added value of EU action

The revision of Regulation (EC) No 223/2009 aims to make the ESS fit for the digital age, strengthening the capacity of statistical offices to make them more responsive in times of crises. By its very nature, the objective of the revision can only be better achieved because of the scale and effects of the proposed action at EU level and clearly cannot be sufficiently achieved by the Member States alone. This is all the truer since the underlying phenomena, digitalisation and the creation of a common digital market, are already within the scope of EU action. Moreover, reuse of privately held data as by-products of digital services could have a cross-border aspect, for instance if the relevant data are kept by multi-national enterprises.

The added value of new, timelier or more granular European statistics at EU level lies primarily in their significance to various policy areas of the Union and their relevance to the Union political priorities (i.e., European Green Deal, an economy that works for people and a Europe fit for the digital age). An analysis carried out by the European Commission in the preparation of the Data Act proposal shows, for example, that a 20% increase in the supply of official statistics would generate an additional EUR 4-12 billion a year in the EU from direct and indirect effects alone[25].

[25] European Commission, impact assessment Report accompanying the Proposal for a Regulation on harmonised rules on fair access to and use of data (Data Act), SWD(2022) 34 final.

4. OBJECTIVES: WHAT IS TO BE ACHIEVED?

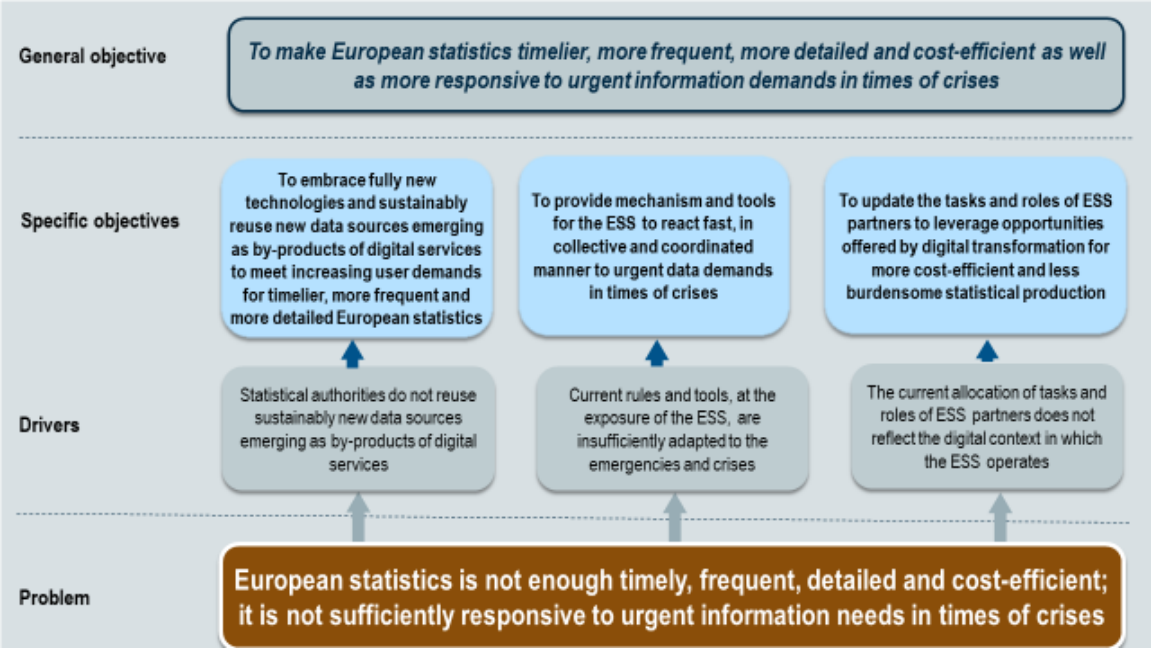

Figure 2: The intervention logic

4.1. General objective

The general objective is to solve the problem as defined in chapter 2, namely to make European statistics timelier, more frequent, more detailed and cost-efficient as well as more responsive to urgent information demands in times of crises.

Realisation of the general objective implies that the difficulties in sustainably accessing and reusing privately held data as by-products of digital services have been overcome, that clear rules and safeguards apply to such reuse throughout the EU, and that demand in times of crises is met. Moreover, the ESS will have improved its efficiency, both in respect of its internal data sharing and task distribution, as well as its effectiveness in providing consistent and accurate measurement of cross-border phenomena. It will exploit the possibilities of the data ecosystems of which the ESS partners are part. The ESS will produce statistics that are more relevant, faster, better, and more detailed. The costs and burdens on the Member States and respondents will have decreased and the ESS will be fit for the digital age and times of crises[26].

4.2. Specific objectives

The overall objective of the initiative will be achieved through three specific objectives.

[26] European Commission (2022), European Statistical System – making it fit for the future, Call for evidence for an impact assessment.

Specific objective 1: To embrace fully new technologies and sustainably reuse new data sources emerging as by-products of digital services to meet increasing user demands for timelier, more frequent and more detailed European statistics[27].

This specific objective is linked to the difficulties in accessing and reusing sustainably privately held data emerging as by-products of digital services for the compilation of European statistics, i.e., the first problem driver. This problem driver was explained in chapter 2 by noting (1) that the EU legal framework, in particular Regulation (EC) No 223/2009, does not provide for a mandate regarding the reuse of such data, (2) that the legal systems of the Member States, with a few exceptions, do not include such a mandate either, and (3) that as a result of the disparities among Member States, no clear rules and safeguards apply to such reuse throughout the EU.

Addressing the difficulties that currently impair the access and reuse of new digital sources especially those that are privately held will allow for embracing fully the opportunities provided by those sources for producing relevant, timely and enough detailed European statistics in both times of crises and in regular times free of unexpected events.

Specific objective 2: To provide a mechanism and tools for the ESS to react fast, in a collective and coordinated manner to urgent data demands in times of crises

This specific objective, which is linked to the second problem driver, involves establishing of a mechanism and toolbox that will ensure that the ESS can react fast and in a collective and coordinated manner during the crises and can provide timely and relevant statistics that are comparable across Member States to support decisions. Within the usual planning cycle, the incubation time for a new statistic typically covers several years, but if the demand is urgent and important, a fast-track solution is needed with appropriate safeguards for all partners and for the quality and harmonisation of the resulting statistical information. This could include producing statistics based on new data collections, calculating new indicators or providing additional insights based on existing data.

The realisation of specific objective 1 would, in particular, facilitate reaching the second specific objective. Only by tapping the new digital data sources can the timeliness of official statistics be substantially improved, in some cases potentially to close to real-time.

Specific objective 3: To update the tasks and roles of ESS partners to leverage opportunities offered by digital transformation for more cost-efficient and less burdensome statistical production

This specific objective is linked to the fact that the ESS is not sufficiently efficient. Limited data sharing results in extra work for the ESS and a higher burden than necessary on businesses and citizens. Since data sharing takes place on the basis of Regulation (EC) No 223/2009, the specific objective implies that barriers need to be removed resulting

[27] Although digital data may have quality issues in some cases, such as their veracity or representativeness, their reuse for official statistics will take place while respecting the quality norms as stipulated in Regulation (EC) No 223/2009.

from diverging EU and national rules and practices. In particular, the aim is to enable data sharing for the production of statistics in areas with a high degree of European integration, such as business trade or cross-border phenomena such as migration and labour mobility.

The current task distribution results in extra work for the ESS and suboptimal collection of data from sources with data from multiple countries. An optimal task distribution could, for instance, entail the possibility for Eurostat to access data from data holders that operate in multiple Member States, on their request.

Measuring cross-border phenomena may require that some statistical actions are performed at the EU level that cannot be done at national level only. This concerns, for example, the compilation of European aggregates or ensuring coherence and consistency of national estimates in domains where the European aggregates cannot be produced as a simple sum of national aggregates. Examples of such domains are globalisation (to avoid double counting of some multinational enterprises and where national statistical authorities can only observe the relevant operations of the enterprises in their country) or global migration flows.

This specific objective will also include outlining possible roles and specifying new functions that statistical authorities could perform in the emerging European and national data ecosystems in full respect of the subsidiarity principle, while recognising that not all statistical authorities are yet in a position to perform all functions related to these new roles at national level.

5. WHAT ARE THE AVAILABLE POLICY OPTIONS?

5.1. What is the baseline from which options are assessed?

In the baseline option (PO0) no revision of Regulation (EC) No 223/2009 is foreseen. Non-binding measures that are currently foreseen or being implemented are part of the baseline option. For instance, Eurostat is currently fostering the development of methodological and quality frameworks for integrating new data sources into European statistics as well as knowledge sharing and acquisition of new skills to enable the use of new technologies and methods in the ESS.

Apart from current and foreseen measures, the changing data environment of the ESS is also relevant to the baseline. The European Data Strategy is profoundly changing the environment of the ESS, affecting, for instance, the availability of open data, accessibility for statistical purposes of non-open data, standardisation and interoperability, data market rules, data governance rules, and possibly also public and business attitudes towards data sharing for statistical purposes. Particularly relevant to the ESS in this context are the following acts and initiatives[28]:

- The Open Data Directive and its implementing acts. These will result in wider availability of open data, in particular in the form of 'High Value Datasets'.

[28] For more information and references see Annex 5.

- The Data Governance Act. The 'common European data spaces' that are being created in this context are potentially very relevant to the ESS, to the extent that data become available for reuse for European statistics. The work of the European Data Innovation Board, created through the DGA, will enhance, among other things, data and metadata standardisation, and interoperability, which is obviously important to European statistics. The DGA also regulates so-called data altruism, with potential benefits to European statistics.
- The Data Act proposal. For cases of emergencies and other exceptional needs, and only for those cases, this act may result in limited and strictly conditional data access for official statistics.
- The Single Market Emergency Initiative. Data collected by the Commission on the basis of the SMEI can be used, under certain conditions, for European statistics. The data would refer to targeted information from the economic operators in crisis-relevant supply chains.
- The Interoperable Europe Act proposal. To the extent that this act results in better cross-border and public sector interoperability, and actual data sharing through the 'Interoperability Europe Portal', this may benefit the ESS.

It is worth noting that these acts and initiatives make the baseline dynamic, which is important because it is used as benchmark for assessing the impact of the other policy options.

5.2. Description of the policy options

Apart from the baseline option, two policy options were designed to achieve the general objective of this initiative, namely to make European statistics timelier, more detailed and cost-efficient, as well as more responsive to urgent information needs in times of crises. Each policy option contains measures that are classified in three groups, each of which corresponds to a specific objective. In this way, it is ensured that all three specific objectives will be well covered by the planned measures.

Policy option 0, the baseline option, implies no revision of Regulation (EC) No 223/2009. This option focuses on preserving the existing incentives and mechanisms to nudge the voluntary access and reuse of emerging digital sources that are privately held for the compilation of European statistics, voluntary participation in collective and coordinated statistical activities in times of crises and non-binding measures to promote more efficiency and less burden on businesses and Member States' national statistical authorities. The measures under this option involve minimal intervention and are non-legislative by nature.

Both policy options 1 and 2 include a targeted revision of the Regulation (EC) No 223/2009 but they differ in the intensity and stringency of the legislative measures (amendments) on the businesses and Member States' national statistical authorities. **Policy option 2** is more ambitious in terms of imposing obligations on the data holders and on the Member States' national statistical authorities whereas **policy option 1** is less stringent with lower intensity. The measures covered by this option aim at introducing legal certainty and empowering the key actors (the data holders and the ESS members) to harness opportunities offered by the digital age and to have adequate instruments for responding fast and in a coordinated way to urgent user demands in times of crises. Data reuse arrangements foreseen in policy option 1 should enable the reuse of digital data for

compiling European statistics in a framed way, while reducing the burden on businesses and Member States' national statistical authorities.

A summary of the main features of the three policy options is provided in Table 1.

The three policy options have the following in common:

- Each option builds on earlier analyses and discussions with stakeholders;
- Each option assumes compliance with applicable rules of data protection, trade secrets and statistical confidentiality;
- Each option foresees transparency measures.

Table 1: Policy options and measures

Policy Option 0	**Policy Option 1**	**Policy Option 2**
• ***The baseline option*** • ***No revision of Regulation 223/2009*** • ***Non-binding measures*** • ***Leverage the general provisions of other EU legal acts*** • ***Continuation of existing measures***	• ***Revision of Regulation 223/2009*** • ***Low legal intensity of amendments*** • ***Less stringent obligations on businesses especially on small and micro enterprises and on national statistical authorities*** • ***Focus on legal certainty and empowering data holders and national statistical authorities***	• ***Revision of Regulation 223/2009*** • ***High legal intensity of amendments*** • ***More stringent obligations on businesses and national statistical authorities***
Specific objective 1 **To embrace fully new technologies and sustainably reuse new data sources emerging as by-products of digital services to meet increasing user demands for timelier, more frequent and more detailed European statistics**		
Policy Option 0	**Policy Option 1**	**Policy Option 2**
1.1: Recommendations, exchange of best practices, grants aimed at enhanced re-use of new data sources: • Create financial and non-financial incentives and mechanisms (e.g., ensuring the public recognition of involved stakeholders, or creating incentives for NSIs to use innovative techniques) to increase the re-reuse of new and innovative data sources in the production of European statistics.		

• Develop guidelines based on best practices for establishing partnerships and voluntary data agreement models.		
1.2: Develop common standards on data quality in the context of B2G4S data sharing for the purpose of European official statistics.	**2.1**: *Same as measure 1.2.*	**3.1**: *Same as measure 1.2.*
	2.2: Provide transparency obligations for both private and public actors engaging in data sharing collaboration.	**3.2**: *Same as measure 2.2.*
	2.3: Introduce enforceable mechanisms for reusing emerging new digital data sources that are privately held, for the compilation of European statistics, under specific conditions and subject to a set of binding safeguards: • The right of having access and reuse data includes a consultation stage where the feasibility of various parameters of the data requests (e.g., level of aggregation, deadlines, mode of data provision, confidentiality protection) are discussed. • Compensation limited to costs related to data processing and data extraction needed to make the data usable for the compilation of European statistics. • Dispute resolution mechanisms are foreseen. • Exemption of micro and small enterprises from enforceable mechanisms for reusing emerging new digital data sources that are privately held, for the compilation of	**3.3**: *Measure with higher legal intensity than 2.3:* Introduce enforceable mechanisms for reusing emerging new digital data sources that are privately held, for the compilation of European statistics: • Obligation to appoint data stewards to facilitate collaboration with statistical offices in data re-using activities. • No consultation stage is included on the various parameters of data requests. • No compensations of any costs incurred by the data holders. • Dispute resolution mechanisms are not foreseen. • No exemption of micro and small enterprises from enforceable mechanisms for reusing emerging new digital data sources that are privately held.

	European statistics.	
	2.4: Limit the access and reuse of privately held data to NSIs and Eurostat, only for the compilation of European statistics.	**3.4**: *Measure with higher legal intensity than 2.4:* • Extend the right to access and re-use privately held data to the other national authorities that produce European Statistics. • Allow the data made available by private data holders to be shared with the research community.
Specific objective 2 **To provide mechanism and tools for the ESS to react fast, in collective and coordinated manner to urgent data demands in times of crises**		
Policy Option 0	**Policy Option 1**	**Policy Option 2**
1.3: Recommendations, exchange of best practices, financial support to enable the ESS to react fast in times of crises: • Develop recommendations to NSIs for preparedness and resilience action plans. • Using financial support to stimulate the ESS partners to participate in voluntary statistical actions in times of crises.	**2.5**: *Same as measure 1.3.*	
	2.6: Provide a legal basis for Eurostat to initiate statistical actions conducted at EU level in response to urgent user demands in times of crises, with voluntary participation of NSIs, but ensuring that it will result in sufficiently timely, frequent and detailed representative data and information at EU level. Such statistical actions could only be initiated in the case a crisis mechanism has been formally triggered by an Institution according to established procedures	**3.5**: *Measure with higher legal intensity than 2.6:* Provide a legal basis for Eurostat to initiate statistical actions conducted at EU level in response to urgent user demands in times of crises, with obligation of the Member States to take part in the actions, ensuring that it will result in sufficiently timely, frequent and detailed representative data and information at EU level.

	in Union law; it would be not for the ESS to decide on its own on the existence of a crisis.	
		3.6: Oblige the Member States (NSIs) to establish resilience mechanisms and crises preparedness action plans to ensure that national statistical systems will be able to function in times of crises and respond to urgent users demands.
		3.7: Provide a legal basis for Eurostat to initiate statistical actions conducted at EU level, with voluntary participation of NSIs, in response to urgent user demands, other than in times of crises.
Specific objective 3 **To update the tasks and roles of ESS partners to leverage opportunities offered by digital transformation for more cost-efficient and less burdensome statistical production**		
Policy Option 0	**Policy Option 1**	**Policy Option 2**
1.4: Recommendations, exchange of best practices, communication activities, grants, in particular: • Develop common standards regarding technical interoperability of data across borders and sectors for the ESS. • Develop guidelines and recommendations for voluntary data sharing within the ESS. • Support projects aimed at enhancing data sharing within the ESS through existing collaboration networks defined in Regulation (EC) No 223/2009.		

• Develop recommendations for more active participation of NSIs in the emerging data ecosystems at national level.		
	2.7: Empower Eurostat to collect data on behalf of the ESS and act as a data hub to share data with NSIs for reasons of efficiency or effectiveness, in particular in domains related to cross-border flows and phenomena, and in cases where this would reduce the burden on businesses and citizens or the workload of the ESS.	**3.8**: *Same as measure 2.7.*
	2.8: Establish that NSIs and Eurostat may assume data governance and data stewardship functions, for instance in respect of standards and data interoperability within their data ecosystems.	**3.9**: *Same as measure 2.8.*
	2.9: Foresee possibility for Eurostat to initiate the development of experimental statistics in close cooperation with the ESSC.	**3.10**: *Same as measure 2.9.*
	2.10: Make data sharing among NSIs and between NSIs and Eurostat mandatory for statistical purposes based on cost-benefit analysis, in particular in domains related to cross-border flows and phenomena and: • Empower NSIs and Eurostat to share these data for statistical research purposes. • Create support mechanisms and incentives to move towards data sharing based on privacy-enhancing technologies (PET).	**3.11**: *Measure with higher legal intensity than 2.10:* Make data sharing among NSIs and between NSIs and Eurostat mandatory for all statistical domains and: • Empower NSIs and Eurostat to share these data for statistical research purposes. • Mandate applications based on privacy-enhancing technologies (PET) to be used for the data sharing within the ESS.

5.3. Options discarded at an early stage

No options were discarded at an early stage. The three options were fully assessed.

6. WHAT ARE THE IMPACTS OF THE POLICY OPTIONS?

6.1. Introduction

This chapter provides an assessment of the policy options in terms of their impacts. These impacts are described qualitatively and, to the extent possible, quantitatively, and where relevant with differentiation between groups of stakeholders. This assessment includes an impact analysis of PO0, the baseline option, because this is the benchmark for assessing the impact of PO1 and PO2.

The impact of the three policy options have been assessed against criteria of effectiveness (in terms of achieving the specific objectives), efficiency (benefits and costs in reaching the specific objectives), coherence (with other policy and legal initiatives) and feasibility (technical and non-technical, taking into account the political context). Given the domain of the initiative, i.e., European statistics, specific attention has also been given to the potential of policy options to reduce the burden on businesses and respondents as well as to the distributional effects of the impact (e.g. the effect on small and micro enterprises will be different from the effect on large businesses holding data). The impact analysis takes also into account the risks that benefits that cannot be realised because of, for instance, issues with the quality of the new data sources.

In the current initiative, the stakeholder groups that are most relevant to distinguish in respect of impact include (see also Annex 2):

- providers of primary data for the production of European statistics (including data holding businesses, individual respondents, and public administrations);
- producers of European statistics, i.e., the partners of the ESS (in particular the NSIs and Eurostat);
- users of European statistics (including institutional and business users, the media and the general public);
- other stakeholders, including subjects to which data collected for European statistics pertain.

It is noteworthy that businesses appear in this grouping in two roles: as data holders and as users of European statistics. For businesses, moreover, a distinction in respect of size is also very relevant, in particular distinguishing small and micro enterprises. Similarly, citizens may be subjects of data collection, but are also part of society at large, whose functioning depends on the public availability of impartial and high quality official statistics. Hence, the social impact of the initiative is also considered in the impact analysis.

For the analysis of the impact of the policy options, several sources of evidence have been used (see also Annexes 1 and 4). Where possible, a distinction has been made between direct, indirect, one-off and recurrent costs and benefits. The quantification of impacts, where possible, was based actual experiences with B2G4S sharing of specific types of digital data held by the private sector and with internal ESS data sharing (i.e. Intrastat data sharing), to which several measures of PO1 and PO2 refer. Historical data, existing studies and experts' judgements were also heavily used to derive quantitative

estimates of benefits and costs of policy options. In addition, several use cases of B2G4S data sharing (described in Annex 6) have been also used as input for the quantification of costs and benefits related to the various policy options. For the calculation of the quantitative impact of the policy options, assumptions had to be made as to the number and type of use cases that will be realised following the adoption of the revision of Regulation (EC) No 223/2009. When describing the costs and benefits of the policy options in this chapter, the uncertainty of the assumptions is taken into account.

It is important to note that the use cases presented in this report are not meant to determine the scope of access in the revised legal framework. Even if limited in number, they do represent the most typical cases of new data sources with the highest potential to address the problem identified under this initiative and as such provide valuable information in terms of expected impact. These use cases represent therefore a fair basis for the measurement of the direct and indirect benefits as well as of the costs.

The overall direct benefits and costs were estimated using a standard cost-based model that considers direct costs and benefits for modifying existing or setting up new data collections for producing enhanced statistical output. The model distinguishes between costs and benefits incurred by the producers of official statistics and by the respondents (here the data holders). The costs are further distinguished according to their type: upfront (preparation), organisational, infrastructure and operational costs, which are shared between statistical offices and data holders. The considered benefits are quality improvements and increased statistical output from which statistical offices are profiting. Cost savings due to smaller number of respondents are incurred by statistical offices and are mirrored as burden reductions at the side of the businesses. It has to be emphasized that the costs are incurred by the data holders while the savings due to burden reduction are incurred by the entire business sector in general.

An extensive stakeholders' consultation has also fed into the analysis of the impact of the policy options, but these inputs are mostly qualitative (see also Annexes 2 and 4). The dimensions looked at included:

- employment and economic growth
- technological development and the digital economy
- innovation and research
- EU evidence-based policymaking
- conduct of business, administrative costs on businesses and sectoral competitiveness
- position of SMEs
- public authorities and their budgets

When the impact of policy options is considered, it should be kept in mind that the Regulation (EC) No 223/2009 is a framework regulation with an enabling character. The foreseen revision of the Regulation will have its full effect through subsequent decisions, at both the EU and national level. For example, the concrete demands on a specific digital data source and its concrete use for compilation of European statistics will be carried out through the Annual Work Programme for European statistics (AWP) adopted by the Commission and based on conditions specified in the basic act. In accordance with Regulation (EC) No 223/2009, the Commission must ensure, when preparing the AWP, effective priority setting, including reviewing, reporting on statistical priorities and allocation of financial resources; these are important elements to take into account when

adjusting the statistics referred to in the AWP to the evolving needs of the users and eventually adding or removing statistics. The draft AWP lists the eligible actions and refers to the main activities and outputs as well as to the main statistics produced and disseminated by Eurostat. The AWP is submitted to the ESS Committee and the Commission must take the utmost account of the comments of the ESS Committee. Every time a statistic is added to the AWP that could make use of privately held data, either as its main source or as a supplementary source, this will require a convincing justification, including a test of the proportionality of the costs and benefits to society and to the concerned data holders. When executing the AWP, specific data holders will be identified and requested to enable the reuse of the data held by them. Such requests must also be explicitly justified. The actual impacts will thus be evaluated on a case-by-case basis before relevant decisions are taken on the basis of the future revised framework regulation. This is further explained in chapter 8.

The actual impact of the policy options will thus take some years to become fully visible. For the impact analysis of the remainder of this chapter, a timeframe of 10 years is used as reference.

6.2. Impact of policy option 0: the baseline option

The baseline option assumes that the current regulation on European statistics is not changed. Ongoing efforts of the ESS to meet the increasing demand for European statistics would be continued as described in chapter 5. The opportunities provided by new legislative measures including the Open Data Directive, the Data Governance Act, the Data Act proposal, the Single Market Emergency Initiative and the Interoperable Europe Act proposal would be utilized.

Measures under this policy option are of non-binding nature. They rely on voluntary participation supported by limited financial contributions at European level. Therefore, their impact will depend on their take up by the stakeholders. Based on accumulated experience in the last decade, the measures related to the first specific objective (concerning B2G4S data reuse) are expected to result in a very slow uptake of B2G4S data sharing. These measures will lead to participation of only few data holders in data sharing activities at EU level and in some Member States, mainly those who have additional legislation in place to enable data sharing activities. The voluntary character of the actions will most likely not reach a level to replace or substantially decrease the scope of current data collections. The resulting statistics will most likely not meet the high-quality requirements, notably comparability and cost-efficiency, as expressed in the regulation on European statistics due to the risk of non-take up of new data sources by each Member State. This would lead to a more heterogeneous situation and a growing gap among the Member States. It would very likely render impossible to compile comparable European statistics from such a disperse situation. This assessment is confirmed by the results of the stakeholder consultation and by own experiences from projects conducted by EU statistical offices for the use cases described in Annex 6 (mobile network operators data, smart meter, financial transactions, scanner data).

The measures related to the second specific objective (concerning crisis response) and the third one (concerning efficiency) are also expected to have a very limited effect due to their entirely voluntary nature, reflecting current experience. Crisis response would be improvised and not be harmonised, comparable to the insufficiently fast and coordinated

response to the urgent information demands under COVID-19 crisis, and internal ESS data sharing would remain clearly suboptimal.

The baseline option entails the impact of the other existing or under negotiations legislative acts, outside the statistical legislation. The implementation of Common European Data Spaces (CEDS) and activities related to achieving interoperability within and between data spaces could contribute to improving availability of data for the production of European statistics. Considering the principle of voluntary participation in the CEDS[29] and the right of each individual data supplier to determine the type and conditions of use for each case of data reuse, the required quality conditions for the production of European statistics are difficult to meet. Reasons are the insufficient number of participating data holders, the lack of application of common statistical methods and standards, and insufficient transparency in pre-processing of the source data. However, CEDS can be used to further experiment, develop methodologies, quality frameworks or standards for integrating data into European statistics.

The SMEI and the Data Act limit the use of data for statistical purposes to very narrowly defined measures of crisis management and recovery. Hence, it is unlikely that the narrow definition of the SMEI would constitute conditions for producing official statistics that would comply with the quality criteria of European statistics. The conditions in the Data Act are limited to cases of exceptional need.

Despite the uncertainties accompanying the implementation of the baseline scenario, the expected costs and benefits have been estimated, to serve as a benchmark for comparison with the other two policy options. More detailed calculations are presented in Annex 3. The quantification of the impact has been made under the assumption that, in 10 years, PO0 will result in one case of B2G4S reuse of data at the national level (for 18 out of 27 NSIs), one case of effective ESS response to a crisis (new data collection or new statistical insight, e.g., people mobility flows under lockdowns) and one case of data sharing within the ESS (e.g., on data on multinationals). This assumption can be considered realistic, since it is derived from the extensive experience of implementation of measures foreseen to be continued under PO0 and associated risk analysis of non-taking up the measures by the data holders and the ESS partners.

The estimates show that the balance of all direct costs and benefits would be positive and amount to EUR 87.2 million a year (Table 2, C12 and C13), but, importantly, the baseline option would not realise much higher benefits through intensified B2G4S data sharing, mechanisms for responding to urgent user demands, and intensified ESS data sharing. Moreover, already declining response rates would take their toll, and further investments would still be needed. These conclusions would remain valid and would not be affected as such with a slight change in the assumptions concerning the three cases of B2G4S data reuse, crisis response and data sharing. The results of the impact assessment would continue to hold.

Concerning the different categories of stakeholders, additional burden on **data holders** could be generated through the implementation of the above-mentioned legislation, mainly the Data Governance Act. However, the benefits created by these legislative acts

[29] Commission staff working document on Common European Data Spaces, SWD(2022) 45 final.

would outweigh additional burden on private enterprises[30]. Additional legislation related to sectoral European data spaces could induce additional burden on public administrations through data sharing obligations.

The burden on **producers of European statistics** would slightly increase due to some voluntary agreements of data reuse, activities contributing to achieving interoperability, methodological developments, a voluntary participation in statistical actions under crises and data sharing within the ESS.

Users of European statistics would profit only in a limited way, proportionally to the increase in European statistics output, mainly based on more intense use of administrative data. The increase will very unlikely close the gap between the supply and the demand for more granular and timelier statistics. The growing gap will have a negative impact on the quality of public debate and the use of European statistics for evidence-based policy making. The lack of European statistics is more severe in domains heavily affected by digitalisation, such as the digital economy, as traditional methods of statistical data collections are not adequate to provide high-quality information.

The baseline option does not directly change the burden on **SMEs**. However, opportunities to lower the burden on SMEs through alternative ways of data collection are missed in this scenario. In addition, an increase in quality of current statistics through collection of data via larger data holders is only possible to a very limited extent. An example of this effect is the collection of statistical data on accommodation services from internet platforms, which cover the offers from enterprises exempted from traditional data collections due to their size.

The baseline option is unlikely to significantly impact employment and growth, technological development, innovation and research, or the situation of SMEs and public authorities. On the contrary, the baseline scenario would rather have negative effects of EU evidence-based policy making, because of the dynamic nature of the baseline with a increasing gap between demand and supply of statistics.

6.3. Impact of policy option 1: the first legislative option

For PO1, the impact is analysed for the measures associated with each of the specific objectives, respectively. Where an impact is realised by the combination of measures of more than one specific objective, this is indicated.

Overall impact by specific objective

Due to the obligation of making their data available for compilation of European statistics, PO1 will impact data holders more intensively. Concerning **the first specific objective**, the burden on data holders will increase as compared to the baseline option. This will affect those enterprises which are in the possession of large digital datasets suitable for producing European statistics. Examples are enterprises holding data from

[30] European Commission, impact assessment Report Accompanying the document Proposal for a Regulation on European data governance (Data Governance Act), SWD(2020) 295 final, and European Commission, impact assessment Report accompanying the Proposal for a Regulation on harmonised rules on fair access to and use of data (Data Act), SWD(2022) 34 final.

smart meters, financial transactions, metadata from mobile phone communication records and from internet platforms. Usually, these enterprises can be categorized as big enterprises. Due to market concentration, typically few enterprises cover a large part of the statistical target population. In case of mobile communication providers, these are typically 3-4 per Member State; in case of internet platforms, these are between 4 and 600 across Europe depending on the domain (see Annex 3 and the example of accommodation platforms and online job portals of Annex 6). In addition, the largest platforms operate across Europe.

The impact analysis assumes that in the first 10 years after the revision of Regulation (EC) No 223/2009 comes into force there will be reuse of privately held data for European statistics in 15 statistical domains, which are deemed more mature for use of digital data sources based on recent pilots undertaken by the ESS in areas such as labour market, price and social statistics as well as energy and business domains. In some of these cases, data collection will be centralised at EU level, with Eurostat providing a data hub service to the ESS. This role builds on the corresponding measures of specific objective 3. Annex 3 provides a detailed calculation of costs and benefits, taking into account whether such reuse is carried out at central or national level, and whether it is aimed at replacing existing surveys or at compiling new statistics.

European statistics based on B2G4S data sharing require considerable investments in infrastructures and the development of methodologies and processes, with increasing operational costs. On the other hand, savings are incurred due to decreasing samples, and increasing numbers and quality of statistics output. Considering only direct costs and benefits, relevant for the first specific objective within the time span of 10 years it is estimated to achieve net benefits for businesses and the ESS together of approximately EUR 653.5 million a year, of which more than one third would pertain to businesses (Table 2, E3 and E4). Although for businesses the costs would be more than compensated by savings due to burden reduction for current surveys, the distribution among businesses of costs and benefits would be uneven. Data holders would incur costs, whereas SMEs are expected to particularly benefit, since they make up the large majority of respondents in business surveys. This aspect of the impact is given special attention in chapter 8.

In addition to the direct net benefits, there would be considerable efficiency gains through indirect benefits, in particular better policy decisions and, for businesses, efficiency gains due to better informed economic decisions. Society at large would benefit from effects induced by better informed policies and evidence-based public debate. This, however, cannot be quantified.

It should be noted that the costs and benefits of the measures of the first specific objective depend on the assumptions made regarding the number of use cases that will be realised within the time span of 10 years, for the benchmark (PO0) as well as for PO1. Making such assumptions cannot be avoided, given the framework nature of Regulation (EC) No 223/2009. For instance, the estimation of net benefits of EUR 653.5 million a year, based on 15 statistical domains, would be higher if the number of use cases realised was underestimated, or lower if overestimated: the net benefits are roughly proportionate to the number of use cases (see Annex 3). The assumption of 15 cases is based on user demands for statistical data related to different policy areas of the European Union, the maturity of available data sources, the preparatory work of the ESS, which led to a number of pilot applications and development of standardized procedures and the

capacity of the ESS in conducting new statistics production. Thus this assumption is realistic and the conclusions are not changed by a slight variation of use cases. Therefore, the conclusions on the nature of the impact can thus be considered robust.

The measures of **the second specific objective** for PO1 concern mainly the ESS capability to respond adequately to urgent user demands in times of crises. Use would be made of new data sources that emerge as by-products of digital services, thus building on the measures of objective 1. The measures would affect businesses, since they would need to share data.

The impact analysis assumes that there will be five cases of response to urgent user demands in times of crises based on historical experience about urgent statistical demands originating from crises and an assessment of the overall resource capacity of the ESS to deliver new statistical information products. These five cases would all lead to new statistical products. In some of these cases, data collection will be centralised at EU level, with Eurostat providing a data hub service to the ESS. Therefore, this role also builds on the corresponding measures of specific objective 3. Annex 3 provides a detailed calculation of costs and benefits, taking into account whether such reuse is carried out at central or national level.

European statistics based on the capability to respond adequately to urgent user demands in times of crises would also require considerable investments in infrastructures and the development of methodologies and processes, with increasing operational costs. There would also be benefits in the form of increasing numbers and quality of statistics output, but there would not be savings due to lowering response burden. Considering only direct costs and benefits, relevant for the second specific objective**,** within the time span of 10 years it is expected to achieve benefits for the ESS of approximately EUR 10.7 million (mainly due to additional outputs) and a net loss for businesses of approximately EUR 9 million a year (Table 2, E6 and E7). This is the price to be paid for the huge indirect benefits, consisting of being well-informed and take prompt and efficient decisions in times of crises. These benefits cannot be quantified but would pertain to society as a whole.

As was the case for the first specific objective, the quantitative impact of the measures depend on the assumptions regarding the number of use cases. In the case of crisis response, the prediction is based on past experience with the pandemic and energy crisis. Nevertheless, as was the case for the first specific objective, the conclusions on the nature of the impact are fairly robust, since the impact is roughly linear to the number of cases.

The measures of **the third specific objective** for PO1 concern not only the data hub role as mentioned above, but notably also mandatory data sharing among NSIs and between NSIs and Eurostat. In addition, there are measures such as on the possibility for NSIs and Eurostat to assume data governance and data stewardship functions. These additional measures are hard to quantify and are expected to have an impact that is much less than the impact of mandatory ESS data sharing.

The impact analysis assumes that within 10 years there will be four new cases of ESS data sharing. Annex 3 provides a calculation of costs and benefits, building on the experiences with Intrastat statistics. The application of data sharing for the Intrastat statistics resulted in a reduction by EUR 155 million a year, largely eliminating the

collection of mirror statistics. Intrastat had the biggest potential for savings and burden reduction due to the size of the data collection, so this will not be repeated for the four new cases of data sharing. However, a considerable reduction of burden can still be expected when applying this measure to statistical domains such as migration or data collections in business statistics. The impact analysis shows estimated net benefits of EUR 138.3 million a year (Table 2, E9 and E10) for the first specific objective, in a time span of 10 years. These benefits are the result of the burden reduction on respondents that can be realised, together with lower net costs for the statistical authorities. It is primarily the respondents who will benefit.

Expected impact on categories of stakeholders

In addition to the information on the expected impact of the measures on specific groups of stakeholders provided above, some further observations can be made. In total, it can be estimated that around one thousand relatively large **enterprises** across Europe will be affected by mandatory data sharing requests (Annex 3). Considerable savings can be envisaged related to decreasing the size of surveys and thus reduction of burden on businesses and citizens. The reduction of burden will affect all sizes of businesses, including SMEs. The estimations show that the savings due to decrease of samples are more than 10 time higher than the additional burden on enterprises due to new data demands (Annex 3).

Users of European statistics will profit considerably from increased availability of statistical data and of higher quality of European statistics through the application of the measures in PO1. The most effective measure will be reuse of privately held data followed by actions in case of urgent user demands related to crises. Intensified data sharing and Eurostat acting as a data hub will mainly result in savings and burden reduction by increasing the quality and quantity of European statistics. Increased ability of the ESS to provide timelier, more frequent and more detailed statistics in response to urgent information needs in times of crises will contribute to higher impact of initiatives at European level addressing those crises.

Public administrations will profit from increased availability of statistical data and from additional activities of statistical offices related to data governance and stewardship. The ESS and its partners have well-advanced governance mechanisms, processes, methodologies and quality frameworks, and taxonomies, which have the potential to increase interoperability within and between different domains. Although implementation will vary among Member States, all **NSIs** will profit from concepts and guidelines being elaborated by the ESS and by international statistical organisations, such as the statistical divisions of the UN. However, PO1 requires considerable investment from the NSIs in methodology, processes, infrastructure and training of staff due to the very different nature of data collection and data treatment in the context of B2G4S data sharing for European statistics as compared to traditional processes. Nevertheless, necessary investments to realize the potential of the measures of PO1 may be offset by savings. The investments and savings will occur at different points in time, with upfront investment at the beginning of the period required to realize savings at a later stage. Due to the fact that the digital data sources are not optimized for use in official statistics, some extra effort will be necessary to ensure high quality of official statistics outputs.

The **academic sector** will profit from availability of additional microdata for research purposes based on new statistical information products derived from the policy measures.

The current regulation on European statistics already includes the legal framework and limitations for access to microdata for research purposes, which remain fully in place.

6.4. Impact of policy option 2: the second legislative option

For PO2, the impact is also analysed for the measures associated with each of the specific objectives, respectively. Where an impact is realised by the combination of measures of more than one specific objective, this is indicated.

The estimation of direct benefits and costs of the measures related to B2G4S data reuse, crisis response and ESS data sharing follows the same methodological approach as for PO1. For the reasons given in section 6.3, the conclusions on the nature of the impact can be considered robust. The estimation of the number of use cases takes into account different factors related to frequency of crisis based on recent experiences, demand by users for new or improved statistics, capacity of the ESS, readiness of the ESS and the envisaged legal conditions of the policy option.

Overall impact by specific objective

Measures of **the first specific objective** for PO2 also concern the reuse of new data sources that emerge as by-products of digital services, and the measures will impact such data holders more intensively than PO0 and PO1. The burden on data holders will increase as compared to the baseline option. This will affect those enterprises which are in the possession of large digital datasets suitable for producing European statistics. However, the framed reuse conditions, foreseen under PO1 in respect of B2G4S data reuse do not apply to PO2. In particular, there will be no consultation stage to discuss feasibility issues and operational modalities, no dispute resolution mechanisms and no cost compensation including marginal costs related to the preparation of those data for reuse by NSIs and Eurostat. As a result, disputes can only be settled in courts, and sustainable data reuse in the production of official statistics will take much more time to be realised.

As a consequence, to quantify the costs and benefits of PO2, the impact analysis assumes an effective sustainable reuse of privately held data, implemented in 5 European statistics' domains in the first 10 years after the revision of Regulation (EC) No 223/2009 comes into force. In some of these cases, data collection will be centralised at EU level, with Eurostat providing a data hub service to the ESS. This role builds on the corresponding measures of specific objective 3. Annex 3 provides a detailed calculation of costs and benefits, taking into account whether such reuse is carried out at central or national level, and whether it is aimed at replacing existing surveys or at compiling new statistics.

Even though B2G4S data reuse will take more time, European statistics based on such data sharing still require considerable investments in infrastructures and the development of methodologies and processes, with increasing operational costs. On the other hand, savings will be incurred due to decreasing samples and increasing numbers and quality of statistics output. Considering only direct costs and benefits, within the time span of 10 years it is expected to reach net benefits for businesses and the ESS together of EUR 207.8 million a year (Table 2, G3 and G4), of which more than a third would pertain to businesses (Annex 3). Although for businesses the costs would be more than compensated by savings due to burden reduction for current surveys, the distribution among businesses of costs and benefits would be uneven, as was the case for PO1. Data

holders would incur costs, whereas SMEs would particularly benefit, since they make up the majority of respondents in business surveys.

In addition to the direct net benefits, to the extent that B2G4S data reuse is realised in PO2, there would be considerable efficiency gains through indirect benefits, in particular better policy decisions and, for businesses, efficiency gains due to better informed economic decisions. Society at large would benefit from effects induced by better informed policies and evidence-based public debate. This, however, cannot be quantified.

Finally, PO2 also comprises a measure aimed at providing access to the data shared by data holders for research purposes. This will lead to costs for the ESS to provide this access. These costs are estimated at approximately EUR 2 million a year (Table 2, G11 and Annex 3). It is expected that the benefits to society of the research done are much higher, but this cannot be quantified.

The measures of **the second specific objective** for PO2 concern not only the ESS capability to respond adequately to urgent user demands in times of crises. For meeting urgent user demands, use would be made of new data sources that emerge as by-products of digital services, thus building on the measures of objective 1. The measures would affect businesses, since they would need to share data.

The impact analysis assumes that there will be ten cases of response to urgent user demands including in times of crises. In some of these cases, data collection will be centralised at EU level, with Eurostat providing a data hub service to the ESS. Therefore, this role also builds on the corresponding measures of specific objective 3. Annex 3 provides a detailed calculation of costs and benefits, taking into account whether the response is carried out directly at EU level or through coordinated actions at national level.

European statistics based on the capability to respond adequately to urgent user demands in times of crises would also require considerable investments in infrastructures and the development of methodologies and processes, with increasing operational costs. There would also be benefits in the form of increasing numbers and quality of statistics output. Considering only direct costs and benefits, within the time span of 10 years it is expected to reach net benefits of approximately EUR 42.9 million annually for the ESS, mainly due to additional outputs, and a net cost of EUR 2.3 million a year for businesses (Table 2, G6, G7 and Annex 3). The costs for businesses are lower than for PO1 mainly because the scenario includes improvements of statistical products as opposed to producing completely new statistical information, which is the basic assumption for PO1. The indirect benefits, consisting of being well informed and meeting urgent user demands in times of crises, cannot be quantified but would pertain to society as a whole.

As was the case for PO1, the measures of **the third specific objective** for PO2 concern not only the data hub role as mentioned above, but notably also mandatory data sharing among NSIs and between NSIs and Eurostat. In addition, there are measures such as on the possibility for NSIs and Eurostat to assume data governance and data stewardship functions. These additional measures are hard to quantify and are expected to have an impact that is much less than the impact of mandatory ESS data sharing.

The impact analysis assumes that within 10 years there will be six new cases of ESS data sharing rather than the four of PO1, because there will be less conditions attached and based on the analyses of cases of cross-border phenomena where the ESS data sharing

could be realised. Annex 3 provides a calculation of costs and benefits, again building on the experiences with Intrastat statistics. For ESS data sharing, the impact analysis shows estimated net benefits of EUR 176.7 million a year (Table 2, G9 and G10). These benefits are the result of the burden reduction on respondents that can be realised, together with lower net costs for the statistical authorities. It is primarily the respondents who will benefit.

Expected impact on categories of stakeholders

In addition to the information on the expected impact of the measures on specific groups of stakeholders provided above, some further observations can be made. It can be estimated that less **enterprises** than for PO1 will be affected by mandatory data sharing requests due to the lower number of statistical domains. The number of enterprises will be at most a few thousand. Still, considerable savings can be envisaged related to decreasing the size of surveys and thus reduction of burden on businesses and citizens. The reduction of burden will affect all sizes of businesses, including SMEs. However, more enterprises will be affected than for PO1 by requests to provide data in cases of urgent user demands.

Users of European statistics will profit considerably, although less than for PO1, from increased availability of statistical data to higher quality of European statistics through the application of the measures in PO2. The most impactful measures will be reuse of privately held data and actions in case of urgent user demands in times of crises. Intensified data sharing and Eurostat acting as a data hub will mainly result in savings and burden reduction by increasing the quality and quantity of European statistics. The overall societal and economic impact resulting from initiatives at European level in response to urgent user demands will however be much higher.

Public administrations will profit from increased availability of statistical data and from additional activities of statistical offices related to data governance and stewardship. Again, all **NSIs** will profit from concepts and guidelines being elaborated by the ESS and by international statistical organisations, such as the statistical divisions of the UN. As was the case for PO1, PO2 requires considerable investments from the NSIs in methodology, processes, infrastructure and training of staff due to the very different nature of data collection and data treatment in the context of B2G4S data sharing for European statistics as compared to traditional processes. However, necessary investments to realize the potential of the measures of PO2 may be partially offset by savings. The investments and savings will occur at different points in time, with upfront investment at the beginning of the period required to realize savings at a later stage. Due to the fact that the digital data sources are not optimized for use in official statistics, some extra effort will be necessary to ensure high quality of official statistics outputs.

The **academic sector** will profit considerably from availability of additional microdata for research purposes. The current regulation on European statistics already includes certain obligations, but data access will be substantially increased in PO2.

7. HOW DO THE OPTIONS COMPARE?

7.1. Effectiveness

The collected evidence suggests that PO0 would not be effective. The voluntary nature of measures included in the option makes it impossible to address the obstacles of legislative nature to the sustainable use of new and innovative data sources for the production of European statistics. As a result, the existing fragmentation of national legal frameworks, if present, will continue. In such environment, as revealed by the use cases (Annex 6) it will remain difficult for the private data holders to find commonalities in the operational implementation of data sharing for the statistical production.

Regarding the second strategic objective, to equip the ESS with a mechanism and tools to respond fast and in a coordinated manner in times of crises, PO0 will also not be effective. Based on experience with Covid-19 and lessons learnt, the purely voluntary nature of participation of the statistical authorities in activities risks not achieving enough coverage or representativeness of the statistics and statistical insights that would be developed to meet urgent information demands (e.g., if an insufficient number of Member States participate). Even more, those Member States who would be voluntary contributing and investing in statistical actions will be discouraged to do it when the next urgent response would be needed. Achieving a satisfactory level of the preparedness of the ESS for the potential crises will also be strongly dependent on their take up of the recommendations and good practices that will be promoted by the Commission (Eurostat).

Overall, during the consultation process stakeholders made it clear that the general objective of making European statistics more relevant, timelier, more frequent and more granular as well as more responsive in times of crises, could be achieved only if legal action is taken, because all other non-legal possibilities have already been exhausted. To see better statistics, one would need a “quantum leap” that would be implemented only with the legislative measures foreseen in PO1 and PO2. Being not effective in achieving the general goal of the initiative, PO0 would not be effective for closing multiple information gaps (described for example in Annex 7). Taking decision on partial information might lead to suboptimal policy interventions, which can potentially be extremely costly.

While both PO1 and PO2 are more effective than PO0 for all specific objectives, a comparison of PO1 and PO2 shows differences for the three specific objectives. PO1 is by far more effective than PO2 in achieving the benefits of sustainable B2G4S data reuse. This is caused by the conditions for enterprises in PO2. There are less safeguards, no consultation stage, no dispute resolution mechanisms and no cost compensation, which is expected to result in a lower degree of success in B2G4S data reuse, especially given the fact that the ESS will rely on the cooperation of enterprises. Concerning the second specific objective, on the response in crisis situations, the stricter obligations for data holders in PO2 will make that option somewhat more effective, although this depends on the occurrence and the nature of crises. In contrast to PO1, we expect that data production in PO2 would also imply major improvements in existing statistics, which would lower the response burden on enterprises. However, this effect would only be of temporary nature as the crisis measures would be limited in time as well. Similarly, for specific objective 3, on the tasks and roles of ESS partners, PO2 will be somewhat more effective than PO1, due to the stronger internal data sharing obligations. The clear

advantage of PO1 regarding specific objective 1 outweighs the much smaller advantages of PO2 regarding the other two specific objectives. This is true in a qualitative as well as a quantitative sense. It is worth noting that, concerning effectiveness, on the whole PO1 compares favourably to PO2 for the ESS as well as for businesses (see also table 2 of the next section). The increased effectiveness of PO2 would affect the principle of subsidiarity, which has to be taken into account in the overall assessment of the three policy options.

7.2. Efficiency

PO0 would be of limited efficiency because the stakeholders that would be affected would react on their own initiative, based on their assessment of the optimal ratio between results to be achieved and the resources to be spent on that.

The measures under PO1 will entail higher costs than PO0, related mainly to the consultation stage when the parameters of data requests by NSIs and Eurostat (e.g., frequency of data extraction, granularity of data), operational modalities of making the data available, costs implications, privacy and statistical confidentiality protection will be discussed. The costs incurred during the consultation are expected to lead to lower costs during the implementation phase of the contract arrangements between the data holders and the statistical authorities, which means that PO1 is expected to be more cost-effective than P02. In addition, the small and micro enterprises will be excluded from the enforceable mechanism to make the new digital data sources available for official statistics purposes, which can only increase the cost-efficiency of PO1 compared to PO2.

During the consultation, in the online workshop, stakeholders shared the view that PO1 would be the most efficient. According to them, PO2 would be more efficient in answering urgent user needs in times of crises, by reducing the autonomy of the Member States, while PO1 would be more efficient in ESS data sharing making actual applications conditional to cost-benefit analysis. Regarding the sustainable use of new and innovative data sources held by the private sector available for compilation of European statistics, PO2 bears the risk of putting disproportionate costs and burden on data holders.

Stakeholders acknowledged that PO1 and especially PO2 would be more burdensome, but pointed out that the benefits of automated data sharing could reduce the overall burden of current methods of data collections in the medium term. Finally, they underlined that dedicated funding programmes could alleviate concerns for both private data holders and NSIs (in addition to potentially encouraging investments in security and privacy technology). Financial compensation for private data holders could also be a solution. In this regard, PO1 provides more flexibility than PO2.

Table 2 compares the most quantitatively effective measures of PO1 and PO2 with the baseline option in terms of balance of benefits (benefits minus costs of policy measures). While the B2G4S measure goes hand in hand with a reduction of burden, the crisis response focusses on production of new data and statistical insights. The measure of enforced sharing of data within the ESS results in a high burden reduction due to avoiding redundant data collection. Overall, PO1 turns out as the optimal option with considerable reduction of burden that can overcompensate additional burden on data holders. The quantitative analysis therefore suggests that PO1 would be the best option with considerable increase of benefits and burden reduction as compared to PO0. PO2

would result in less benefits for both the statistical and business sector. For the business sector, the relative drop in benefits from PO1 to PO2 is higher than for the statistical sector.

	A	B	C	D	E	F	G
1	**Policy measure**	**Sector**	**Baseline** (Million EUR)	**PO1** (value) (Million EUR)	**Δ PO1** (Million EUR)	**PO2** (value) (Million EUR)	**Δ PO2** (Million EUR)
2	**B2G4S**						
3	Balance of benefits	ESS	24.2	403.3	379.0	150.1	125.9
4		Businesses	21.7	296.2	274.5	103.6	81.9
5	**Crisis response**						
6	Balance of benefits	ESS	0.3	11.0	10.7	43.2	42.9
7		Businesses	-0.2	-9.3	-9.0	-2.5	-2.3
8	**EU data sharing**						
9	Savings	ESS	2.3	24.6	22.3	28.0	25.7
10	Burden reduction	Businesses	39.0	155.0	116.0	190.0	151.0
11	**Microdata for research**	ESS					-2.2
12	**Total**	ESS	**26.8**	**438.8**	**412.0**	**221.3**	**192.3**
13		Businesses	**60.5**	**441.9**	**381.4**	**291.1**	**230.6**

Table 2: Comparison of quantitative impact of policy options

7.3. Coherence

While the non-binding nature of the measures under PO0 makes it very unlikely that this option will be incoherent with the legal context, there is a risk that this option will not be fully coherent with the rapidly changing European context, and not well aligned with the overall goals of EU policies on data and crises response mechanisms.

PO1 is far more coherent and in line with recent EU legislative and policy developments such as the European Data Strategy, the DGA and the Data Act. This was confirmed by stakeholders in the online workshop, who also specified that PO1 would be the most coherent.

The European Data Strategy is based on pillars such as a cross-sectoral governance framework for data access and use, investments in data, strengthening Europe's capabilities and infrastructures for hosting, processing and using data, including interoperability. The foreseen targeted revision of Regulation (EC) No 223/2009 under PO1 will be coherent with the cross-sectoral elements of these legal acts and will specify complementary the conditions, procedures and modalities of data sharing for the purposes of compiling European statistics (B2G4S). The measures under PO2, in particular the lack of a consultation stage, no possibility for compensating any costs including the marginal ones, make this option less coherent with the overall political approach to data sharing taken in initiatives under the European Data Strategy.

Both PO1 and PO2 are coherent with the existing crises response mechanisms and with the proposal for the Single Market Emergency Initiative because the measures will be complementary to those foreseen in these acts and focused only on ways for the ESS partners to be prepared and to provide data in response to urgent demands under the public emergencies defined in these acts.

Finally, the GDPR specifies that further processing of personal data for scientific or historical research purposes or statistical purposes should not be considered incompatible with the initial purposes, in accordance with Article 89(1) of Regulation (EU) 2016/679. As B2G4S would not exclude personal data, these data must be collected, processed and anonymised in full compliance with the GDPR, which is what is currently foreseen in PO1 through several policy measures.

7.4. Feasibility

Non-binding measures foreseen in PO0 are fully feasible. They are not only technically feasible, but they will be supported by all stakeholders. PO2 will create more feasibility issues and is expected to face resistance from both the data holders (on measures to achieve the specific goal 1) and the national statistical authorities (on the measures to achieve the specific objectives 2 and 3). PO1 appears to be easier to be accepted on all three specific objectives. During the consultation phase, stakeholders concluded that PO1 and especially PO2 are likely to have a stronger impact on the conduct of business, administrative costs, as well as public authorities and their budgets, compared to PO0. At the same time, they also highlighted the missed relevant benefit for them (if mandatory data sharing was systematically implemented), which would be the (partial) replacement of direct data collection mechanisms (i.e., surveys), which would alleviate the burden on businesses and households. PO2 puts more burden on businesses and members of the ESS than PO1. Therefore, the stakeholders especially from the private sector expressed clear preference for measures foreseen under PO1.

The support of stakeholders affects the technical feasibility of the measures. This is expected to be especially the case for the first specific objective if PO2 is implemented. B2G4S data reuse depends on the technical cooperation of enterprises, which are expected to be less than enthusiastic under the obligations of PO2, since that option provides less safeguards, no consultation stage, no dispute resolution mechanisms and no cost compensation. The technical feasibility of the second and third specific objective is also not enhanced by measures in PO2 that reduce the autonomy of the NSIs.

Having safeguards against excessive burden and costs are especially important to businesses, but privacy and confidentiality concerns of the wider public also play a role [ICF, section 9.2]. In line with the ‘do no significant harm’ principle, Eurostat has given the effect of the initiative on businesses serious consideration. For achieving specific objective 1, there is no real alternative to creating the possibility of mandatory data sharing for European statistics, given the fact that an exclusively voluntary approach to data sharing has already been seriously tested without resulting in, nor expected to result in, data reuse for statistical purposes at a scale that would respond to the needs of users. The best way to treat potential negative implications on data holders is to give them adequate safeguards (as described for PO1) and to make sure that these are respected during the implementation of PO1, in particular in the context of the Annual Work Programme. The feasibility of this approach is supported by the conclusions of the Expert Group on facilitating the use of new data sources for official statistics [EG

B2G4S], which included independent experts from the business sector. Moreover, it is fully in line with opinions issued by the G20[31], the Conference of European Statisticians[32], the ESS Committee[33], the European Statistical Advisory Committee[34] and the European Statistical Governance Advisory Board[35].

A summary of the assessment of the three policy options against criteria of their effectiveness, efficiency, coherence and feasibility, is presented in Table 3.

PO	Effectiveness	Efficiency	Coherence	Feasibility
0	--	O	O	++
1	++	++	++	+
2	+	+	+	-

Table 3: Comparison of policy options[36]

8. PREFERRED OPTION

8.1. The preferred option and its implementation

The choice of the preferred option

The comparative assessment of the three policy options presented in the previous chapter showed that the baseline option (PO0) has the least desirable outcome in terms of effectiveness, efficiency, and coherence, and the first policy option (PO1) the most desirable, with the second policy option in between. PO0 is the most feasible of the options, but the baseline option is evidently ineffective in achieving the general and specific objectives of the initiative. This clearly points to choosing PO1 as the preferred option.

Also at the level of the individual measures of the policy options, PO1 generally scored better than PO2. However, there are two measures of PO2 that scored somewhat better on the dimension of effectiveness in comparison to PO1 (see section 7.1). These are the measures 3.5 and 3.11 of Table 1 of chapter 5, concerning the obligation for Member States to take part in actions at the time of crises, and the obligations of NSIs concerning data sharing, respectively. These two measures are not part of PO1, not only because they score worse in respect of coherence and feasibility, but above all because they reduce the autonomy of the Member States. Since these measures are not indispensable to achieve the general and specific objectives of the revision of Regulation (EC) No 223/2009, they have not been included in PO1.

[31] The Data Gaps Initiative.
[32] CES declaration on new data sources.
[33] The ESS position paper on access to privately held data which are of public interest and the ESS position paper on the future Data Act proposal.
[34] ESAC Doc. 2022/1, The Data Act proposal and the use of private data for official statistics.
[35] ESGAB annual report 2021.
[36] The scores reflect the expected magnitude of impact : (++) strong positive, (+) positive, (O) no impact, (-) negative impact, (-) strong negative impact

Implementation aspects are discussed next, followed by a closer look at the distributional aspects of PO1, in particular on SMEs.

The implementation of the preferred option

Regulation (EC) No 223/2009 is a framework regulation with enabling clauses and with main effects arising through subsequent decisions at both the EU and national level. Concerning mandatory reuse of privately held data, the list of safeguards and conditions will similarly only take effect through implementing regulations that may, for instance, provide details on the way to ensure respect for business interests, on modalities of reuse, or on time-bound dispute resolution.

Once the revised Regulation (EC) No 223/2009 takes effect, actual mandatory data reuse requires taking two steps. As **the first step**, for any statistic to be partly or completely based on new data sources, a decision has to be taken to add the statistic to the AWP, the process of which was described in section 6.1. As a consequence of the conditions listed under the measures for the first policy option, the proportionality of the addition has to be demonstrated. That is, it will be necessary to demonstrate and document that the benefits for society as a whole can be expected to significantly surpass the costs to society as a whole. For this to take place, an adequate process and criteria have to be worked out.

The **second step** consists of the selection and contacting of the specific data holders that will be requested to enable the sharing of data they hold. When a specific data holder is requested to enable data reuse for such a statistic, this request must be explicitly justified. It is the second step that must incorporate the safeguards and conditions of PO1 for the businesses concerned, including procedural guarantees. This shall include procedures to decide on the modalities of reuse, and mechanisms to resolve disputes between data holders and statistics producers, in particular for cases where no agreement can be reached on the modalities of reuse or the respect of business interests. Throughout the process, transparency, confidentiality and other already existing professional statistical standards and requirements will apply.

Distributional aspects will play a prominent role in the second step. NSIs are not required to pay for the data from the data holders, in line with the rules in place today for accessing and using traditional data sources in European statistics. However, sizeable initial investments or marginal costs may be incurred by the data holders when processing the data, especially for aggregating or running algorithms on the primary data to make them ready for use for European statistics. Possible incentives could also take the form of aggregate customised information derived from the data by the NSIs and provided in return to the data holders.

At the implementation phase, care has to be taken that data reuse is effective. There are risks concerning the quality and usability of the data considered for reuse, and there may be a lack of available skills to deal with new data sources. It is the ESS that has to make sure that data reuse only takes place if professional quality requirements can be met. Likewise, the ESS has to ensure adequate training for its staff and the availability of the skills required on new and innovative data sources for the production of timelier, more frequent and more detailed statistics.

The ESS also needs to be able to develop effective coordinated statistical actions that provide fast and relevant ESS responses to information needs arising in times of crisis. However, the aim is clearly for the ESS to be responsive to information needs arising

from a crisis situation that has been officially declared by an Institution but not to decide on its own on the existence of a crisis. It is only in cases where an emergency mechanism has been formally triggered in accordance with procedures established by Union law that Eurostat should have the capacity to organise in parallel a response at ESS level to meet urgent information needs arising from that crisis and when these needs have not already been addressed through the relevant emergency mechanism. For instance, this could be the case when the Single Market emergency mode has been activated by the Council and that the objective is to meet information needs that cannot be covered by information requests addressed to representative organisations or economic operators in crisis-relevant supply chains. Another example is the need to respond to information needs arising from an energy crisis, for example in the context of a Union alert declared by the Commission when there is a substantial risk of a severe gas supply shortage or an exceptionally high demand of gas occurs. Responding to urgent information needs stemming from a crisis in the field of migration is also an example where a coordinated action by the ESS would prove necessary, for instance in the context of a mass influx of displaced persons from third countries as decided by the Council.

Furthermore, intensified ESS data sharing should take place for the development and production of European statistics and for improving its quality, notably with regards to cross-border phenomena while limiting or reducing the burden on respondents, in full respect of confidentiality.

Although the implementation would have a timeline of several years, the revision would have a tangible effect already after a few years. The timeline assumed for the impact analysis was 10 years for having a substantial number of statistics added to the AWP and integrated into statistical production, and for having a substantial increase in reuse of new data sources. There are no circumstances known that this stage that could prevent Member States to apply Regulation (EC) No 223/2009 directly after its revision. The mechanism for collective and coordinated reaction in times of crises will not require any transitional implementation time. As for the update of the tasks and roles of ESS partners regarding data ecosystems, the optimal use of the opportunities of the evolving data ecosystems is an ongoing process for which only targets can be specified[37].

Small and microenterprises

Regarding data reuse, the safeguards and processes that will be put in place, such as on proportionality, are expected generally not to result in requests for the reuse of data from SMEs. The impact analysis of chapter 6 and the calculations of Annex 3.2 not only show that the population of interest consists of large enterprises, but also that it is only a small part of them that are expected to be subject to data reuse requests. Nevertheless, there may be cases where a medium-sized enterprise plays an important or even dominant data holding role in a specific statistical domain, meriting its inclusion in respect of data sharing for statistical purposes. It is conceivable that there are small or even micro enterprises whose data held are of interest for European statistics, since a small enterprise may hold the data on behalf of a large enterprise, but such cases would be exceptional.

Whereas the costs of enabling data sharing would be incurred by large and, to a much lesser extent, medium-sized enterprises, the benefits to enterprises in the form of lower

[37] Cf, Table 5 on monitoring of impacts in chapter 9.

response burden if surveys can be replaced by new data sources would predominantly be enjoyed by SMEs. Therefore, the overall picture of this initiative is very favourable to SMEs, especially small and micro enterprises.

Nevertheless, micro and small enterprises will not incur costs due to mandatory data sharing. Therefore, a threshold in terms of the size of businesses will apply to mandatory data sharing. For surveys the decision on thresholds for businesses as data providers is taken in the context of the AWP on the basis of its effect on the contents and quality of the statistics concerned and the associated public benefits. However, for holders of private data as by-products of digital services, a blanket exemption for micro and small enterprises is included in the preferred option, irrespective of merits. This is justified, since it provides the strongest safeguard conceivable.

8.2. Estimated impact of the preferred option

The estimations show that the preferred option could bring additional total direct annual benefits amount to approximately EUR 793.5 million a year (Table 2, E12 and E13) as compared to the baseline option. It is assumed that the number of applications using B2G4S data sharing and crisis response will be introduced gradually and will reach an estimated total of 20 statistical domains (use cases) within a time period of 10 years after entry into force of the amended regulation.

The additional total costs for the statistical system under the assumption of implementing 20 data collections are EUR 560.3 million a year (Table 11, J10), broken down to EUR 10.9 million at European level (Table 11, D10 and H10) and EUR 549.4 million at national level (Table 11, B10 and F10). The additional total costs for the data holders are estimated at EUR 195.5 million annually (Table 11, K10). The costs have to be contrasted with the benefits, which are for the business sector mainly related to burden reduction and to some compensations. The additional total benefits for the business sector sum up to EUR 460.9 million a year (Table 11, K15), not including savings through burden reduction due to intensified ESS data sharing of EUR 116 million (Table 12). The total direct benefits for the statistical system are EUR 950 million a year (Table 11, J15).

The costs for European actions are considerably lower as compared to implementation of data collections in all EU Member States. The burden on data holders will also be lower as most of the activities assume to produce new statistics and the data holders will share data only with one party in a harmonised format and do not have to adjust their systems to 27 different requirements. The measure for direct data collection at European level with sharing between all ESS members could therefore be a preferred option in case of European wide acting data holders.

Through increase in data sharing of statistical domains collecting data on cross-border phenomena, experience shows that the response burden on businesses can be at least halved for a specific data collection. In the specific example, the response burden could be additionally lowered by EUR 116 million annually (Table 12, B3). Likely, the savings for the statistical system was 14% totalling EUR 6.5 million for the specific data collection. It is expected to realize a similar approach for other statistical domains. Assuming four similar cases with however lower impact, as the Intratrade represents the largest European statistical data collection, this would result in an additional burden reduction of EUR 116 million and additional annual net savings for the statistical system

of EUR 23 million (Table 12, B4). The additional central annual operating cost are estimated at EUR 0.7 million (Table 12, B5), which would reduce the annual net savings to EUR 22.3 million.

8.3. REFIT (simplification and improved efficiency)

The preferred option is expected to generate scope for burden reduction. The wider and sustainable reuse of emerging new digital data source that are privately held, the possibility to allow Eurostat to act as a data hub, and the mandatory data sharing for statistical purposes and on certain conditions will open room for decreasing sample sizes in some surveys or completely replace them. Data sharing within the ESS will also contribute to the reduction of administrative burden for businesses, citizens and the producers of European statistics. It should be also stressed that micro and small enterprises would be exempted from enforceable mechanisms for reusing privately held data.

The nature of the measures proposed (i.e., the fact that most measures would place additional requirements and therefore burden on certain selected private data holders and entail initial investments for both data holders and the producers of European statistics), made it challenging to quantify the net effects. Therefore provided estimates have to be taken with certain element of caution.

For costs savings, the Table 4 below outlines the expectations regarding the different reuse and data sharing scenarios. By intensifying and facilitating reuse of new digital sources and data sharing, the initiative should reduce burden mainly as a result of a global reduction in the number of surveys and automated and simplified processes. This will be to the benefit of both the public and the private sector, even if selected big data holders would experience an increased burden in terms of providing data. However, this increased burden may also be outweighed by the reduction in response cost in terms of fewer surveys. In addition, the increased effectiveness, brought about by the central data collection relating to cross-border flows and phenomena would reduce the burden on NSIs and internationally operating businesses, and mandatory data sharing on certain conditions would increase efficiency by means of a collect-only-once approach, that would entail a net burden reduction on businesses, citizens and NSIs.

Table 4: Expected REFIT Cost Savings under the preferred option PO1

Description – specific objectives	*Expected REFIT Cost saving*
To embrace fully new technologies and sustainably reuse new data sources emerging as by-products of digital services to meet increasing user demands for timelier, more frequent and more detailed European statistics	EUR 826 Mill[38]

[38] Cf. Annex 3, Table 11: PO1: Cost and benefit for data reuse and crisis response (B12-E12).

To provide mechanisms and tools for the ESS to react fast in a collective and coordinated manner to urgent data demands in times of crises	EUR 0 Mill[39]
To update the tasks and roles of ESS partners to leverage opportunities offered by digital transformations for more cost-efficient and less burdensome statistical production	EUR 139 Mill[40]

8.4. Application of the 'one in, one out' approach

8.4.1. Potential new burden on citizens

Regarding **new burden on citizens**, it should be stressed that Regulation (EC) No 223/2009 is a framework regulation for which reason no immediate direct burden on citizens will flow from the revision. Nevertheless, PO1 entails a potential burden reduction for citizens since the extended use of new data sources is expected to result in fewer surveys. Therefore, the burden reduction for citizens will be mainly measured by time saved on responding to surveys. In conclusion the preferred option will not generate any net burden increase relevant for OI-OO.

8.4.2. Potential new costs on businesses

Regarding **new costs on businesses**, and as pointed out in Annex 3, businesses have a double role as both respondents and users. In their role as users, they will only benefit from PO1. As respondents, they will also benefit from the decrease of response burden as result of the automated data requests and the one-stop shop reporting, that the data sharing of PO1 will bring about. The estimated benefits for businesses amount to EUR 460.9 million, as indicated in Table 11 (K15) [Table 11: PO1: Estimated differential cost and benefit for data reuse and crisis response].

Nevertheless, the use of new data sources will result in costs on a small portion of businesses disposing of large data assets, such as Mobile Network Operator (MNOs), banks, energy providers, or internet platforms. As also indicated in Annex 3, such businesses might incur additional costs or might have certain additional obligations in order to enable statistical authorities accessing their data. The costs for those businesses are estimated at EUR 265.4 million (Table 11, K16) but will be compensated over time by the recurrent offsetting excess cost saving mentioned above. In addition, the compensation for the marginal costs related to making the data available for statistical purposes could be envisaged.

9. HOW WILL ACTUAL IMPACTS BE MONITORED AND EVALUATED?

In general, the existing monitoring and evaluation tools, in place and valid for the statistical production and dissemination of European statistics, will be used. These should

[39] Cf. Annex 3, Table 11: PO1: Cost and benefit for data reuse and crisis response (G12, I12).
[40] Cf. Annex 3, Table 12: PO1: Estimated differential benefits and costs of the measure on mandatory ESS data sharing (B3,B4).

enable an analysis of the effectiveness and efficiency of the suggested statistical initiative and of the quality of the data produced. These tools are:

- The European Statistical Programmes (currently based on Regulation (EU) 2021/690) foresee systematically interim and final evaluation reports[41]. Regulation (EC) No 223/2009, as the framework regulation for European statistics, is indirectly part of this reporting mechanism and will be covered.
- The Eurostat Strategic Plans foresees the follow up of key performance indicators, which also apply to Regulation (EC) No 223/2009[42].
- User satisfaction surveys are carried out on a regular basis[43]

The application of the revised Regulation (EC) No 223/2009 will be monitored and evaluated against the general and specific objectives described in section 4.2. An assessment of their impact will be included in Commission progress report and final evaluation report on the European statistical programme to the European Parliament and the Council in accordance with Article 13(5) of Regulation (EC) No 223/2009.

Eurostat and the NSIs will further improve the standard metadata and quality reporting system for European statistics. This will allow a more sophisticated monitoring and evaluation of the statistical processes used in Member states and of the output disseminated. For example, more detailed information will be available on the use of private and administrative data sources (leading to burden saving) by Member States or on the use of shared services or IT tools (leading to cost savings).

Measuring the progress towards achieving the objectives of the initiative the following list of monitoring indicators based on SMART principles has been defined. The progress on these indicators will be measured against the benchmark targets indicated in the second last column of Table 5. The table also indicates the sources of information that will be used measure the indicators

Table 5: Monitoring of impacts

Operational objectives	**Indicators**	**Targets**	**Sources of information**
Specific Objective: ***To embrace fully new technologies and sustainably reuse new data sources emerging as by-products of digital services to meet increasing user demands for timelier, more frequent and more detailed European statistics***			
Significant	Number of statistics	In 10 years, the	The AWP itself and

[41] Regulation (EU) 2021/690 of the European Parliament and of the Council of 28 April 2021 establishing a programme for the internal market, competitiveness of enterprises, including small and medium-sized enterprises, the area of plants, animals, food and feed, and European statistics (Single Market Programme) and repealing Regulations (EU) No 99/2013, (EU) No 1287/2013, (EU) No 254/2014 and (EU) No 652/2014, OJ L 153, 3.5.2021, p. 1.

[42] The result indicators are the following: User trust in European statistics, Share of users not satisfied with the quality of data and services provided by Eurostat, Statistical coverage, Timeliness of statistics: news releases, Number of new experimental statistics datasets published.

[43] Cf. Evaluation - Eurostat (europa.eu)

increase in statistics of the AWP that are based, partially or completely, on data innovative data sources emerging as by-products of digital services.	that are based, partially or completely, on innovative data sources.	decision has been taken to add at least 15 such statistics to the AWP.	documents produced in the process establishing the AWP. These documents also allow a qualitative evaluation in terms of statistical content (better, more timely, more detailed statistics, statistics on new phenomena, statistics linked to new policy needs).
Substantial increase in the use of statistics that are based, partially or completely, on innovative data sources emerging as by-products of digital services.	Percentage of downloads of statistical products from the Eurostat database of statistics that are based, partially or completely, on innovative data sources.	In 10 years, the downloads of statistical products that are based, partially or completely, on innovative data sources has risen to at least 30%.	The ESS has already a system in place that systematically follows what and how many statistical products are downloaded.
Full application of all required safeguards.	For each statistic of the AWP that makes use of privately held data, documentation is available on prescribed due process, including evidence on proportionality and expected effects on the data holders.	100% compliance	The ESS has a quality monitoring system in place, based on the Code of Practice, including peer reviews and oversight. The monitoring system includes the application of all required safeguards. The Code of Practice will be updated to reflect all needs entailed by the revision of Regulation (EC) No 223/2009.
Specific Objective: ***Mechanisms and tools for the ESS to react fast, in collective and coordinated manner to urgent data demands in times of crises***			
Prompt action by the ESS in response to urgent user demands.	Number of actions undertaken with specific reference to the relevant provisions of the revised Regulation (EC) No 223/2009.	In 10 years, the decision has been taken to invoke the relevant provisions at least 5 times.	The ESS Committee will always be informed and consulted on such actions. The ESS Committee has official and extensive records on the execution of all its responsibilities. The minutes of the ESS Committee will also be one of the sources to qualitatively evaluate the adequacy of the actions undertaken.

Specific Objective: ***Update of the tasks and roles of ESS partners to leverage opportunities offered by digital transformation for more cost-efficient and less burdensome statistical production***			
Increase data sharing within the ESS with burden reducing effects.	Increase of data sharing agreements with burden reducing objectives.	In 10 years, at least 4 such additional data sharing agreements.	All data sharing between partners of the ESS is based on agreements that include the justification of data sharing. Where these agreements are currently not systematically monitored, this will be done in the future, starting with the revision of Regulation (EC) No 223/2009 coming into force.
Update the tasks of ESS partners	Uptake of centralised data collection from aggregate data sources (e.g., web scraping agencies) as well as individual data holders (such as EU-wide operating businesses).	In 10 years, at least 5 of the additions of statistics based, partially or completely, on data held by the private sector, are at least partially based on centralised data collection.	For each instance of centralised data collection on behalf of the ESS, the ESS Committee will be consulted. Hence such data collection will be documented and also be available for evaluation.
Realisation of gains in effectiveness regarding cross-border phenomena through data sharing.	Increase of data sharing agreements with effectiveness objectives regarding cross-border phenomena.	In 10 years, at least 4 such additional data sharing agreements.	All data sharing between partners of the ESS is based on agreements that include the justification of data sharing. Where these agreements are currently not systematically monitored, this will be done in the future, starting with the revision of Regulation (EC) No 223/2009 coming into force.
Outline possible new roles of statistical authorities in the data ecosystem	Uptake of new roles in the data ecosystem by ESS partners.	Uptake of new roles in at least two thirds of the Member States, in 10 years, according to peer review reports.	Given the fact that the operational objective is linked to subsidiarity and the indicator is qualitative in nature, the monitoring will take place by the individual partners of the ESS, with reporting in the context of the Code of Practice (e.g., peer reviews).

ANNEX 1: PROCEDURAL INFORMATION

1. LEAD DG, DECIDE PLANNING/CWP REFERENCES

The proposal for a Regulation amending Regulation (EC) No 223/2009 on European statistics was prepared under the lead of Eurostat. In the DECIDE Planning of the European Commission, the process is referred to under item PLAN/2021/11938.

This initiative is not included in the Commission Work Programme.

2. ORGANISATION AND TIMING

An Inter-Service Steering Group (ISSG) assisted Eurostat in the preparation of the impact assessment and legal proposal. It included representatives of Commission services from 25 Directorate-Generals, including the Commission's Legal Service and Secretariat General.

The ISSG contributed to the initiative's preparation in February 2022 (discussion on the consultation strategy, the Call for Evidence and the terms of reference for the support study) and a written consultation in June 2022 on the public online consultation questionnaire. Following a written consultation, one ISSG meeting (6 December 2022) reviewed the draft impact assessment before submission to the Regulatory Scrutiny Board (RSB).

The Call for Evidence was published on 21 February 2022 and was open to feedback from all stakeholders on the Better Regulation Portal for a period of 4 weeks. The public online consultation was launched on 19 July and closed on 25 October 2022.

The draft impact assessment report and all supporting documents were submitted to the RSB on 14 December 2022, in view of a hearing on 18 January 2023.

3. CONSULTATION OF THE RSB

The Regulatory Scrutiny Board reviewed the impact assessment report and gave a negative opinion on 20 January 2023. The impact assessment was revised as follows.

Comments of the RSB	*How and where comments have been addressed*
(B) Summary of findings	
(1) The problem definition appears too narrow and is not supported by the earlier evaluation. It does not present evidence other than limited stakeholder views.	The problem definition has been revised. It is supported, inter alia, by the evaluation of the European Statistical Programme (ESP).
(2) The intervention logic is not established. The objectives are inconsistent with the identified problems and the range and scope of the proposed	The intervention logic has been revised. The general objective is now closely linked to the problem definition. Specific objectives link the problem drivers to the policy options. The measures of the proposed

options are insufficient to address them.	options are directly linked to the specific objectives.
(3) The analysis of impacts is incomplete for all considered options and does not allow for their comparison in the absence of a well-defined baseline scenario.	The baseline scenario has been elaborated in more details to allow more rigorous assessment of the impacts of the other two policy options. The impact of all options has been presented in tabular form for ease in comparison, with the baseline scenario as benchmark.
(C) What to improve	
(1) The report should make clear how the 'evaluate first principle' has been adhered to. It should clarify the problem definition given that the stakeholder consultation synopsis report suggests that the problems related to the objectives identified in the report are wider than described in the problem section. The evidence, from a limited number of stakeholders, that supports the existence and scale of the problems should be corroborated with other types of evidence coming from the evaluation or other sources. In addition, it should be specified whether the problems can only be attributed to the statistical institutions or to businesses that hold electronic data as well.	Based on the revised problem definition, the evaluation of the ESP now plays a central role as shown by its use in chapter 2. In addition, the problem definition builds not only on a wide range of stakeholders views, but also on a report of the Expert Group on facilitating the use of new data sources for official statistics, composed of experts (of various backgrounds, including the business sector) and impact assessments of other statistical regulations. The scope of the problem definition has been adapted accordingly. Future monitoring and evaluation is covered in chapter 9. The current text also clarifies the problem attribution (chapter 2).
(2) The report should clarify the logic of intervention. Either the general objective should be narrowed, or the problems should be identified differently. Once the objectives are consistent with the identified problems, the report should review the range of options and measures they consist of that could remedy the identified problems and achieve the desired objectives.	The intervention logic has been clarified, starting with the problem definition, and is now easier to understand (see Summary of findings (2)). For each of the policy options, measures have been specified (chapter 5). The impact analysis addresses all options and shows to what extent they each contribute to the specific objectives (chapter 6). Together with an assessment of stakeholders' preferences, the comparative analysis (chapter 7) led to the choice of the preferred option (chapter 8).
(3) The report should explain what constitutes a 'crisis situation' as invoked in the problem definition and whether it is a necessary element to trigger an 'agile' response from the ESS.	It is explained (chapter 2) that a mechanism that would allow the ESS to react in times of crises will be complementary to and support the other crisis response mechanisms at EU level once these are

	activated.
(4) The report should include a well-defined dynamic baseline scenario, which cannot be dismissed from the analysis. The dynamic baseline scenario needs to consider the likely developments affecting the ESS such as the impacts of the recently adopted Data Act and the Single Market Emergency Instrument. The baseline scenario should be quantified to the extent this is feasible and used as a reference to assess the impacts of all considered options.	The baseline scenario has been further developed, including its dynamic aspects. Apart from what is mentioned under comment (4), the dynamic aspects discussed include the growing availability of open data and the development of 'common European data spaces' in the context of the Data Governance Act. The impact of all options has been tabulated and compared, with the baseline scenario – quantified to the extent feasible – as benchmark.
(5) As the range of identified problems is potentially wider than currently presented in the report, the corresponding range of options to address the problems should be expanded and should go beyond only addressing the use of privately held data.	See point (2) under (B) and point (2) under (C). All options have been developed at the level of measures, and linked to the three specific objectives, only one of which concerns the use of privately held data (chapter 5).
(6) The report should clarify what type of assessment will be undertaken to justify the inclusion of certain data collections from private actors in the annual statistical work programmes.	A two-step approach is foreseen, in which the decision of inclusion in the annual work programme is the first step. The second step is the decision on individual data requests. Both steps require proper justification. The type of assessment to be undertaken is explained (chapter 8).
(7) For each option, the report should identify and quantify the corresponding costs and benefits, considering their direct, indirect, one-off and recurrent elements. The estimates should be transparently presented to avoid a risk of double counting. The report should be clearer on the distributional impacts, in particular as regards the data owners. To that end, a proper SME test should be conducted. The report should be more specific about the burden reduction potential of the initiative, linked to possible replacement of traditional surveys with collections of digital data. Once the impact analysis is improved, the report should use its results in the comparison of options and justification of the preferred option(s).	In the previous version of the report the costs and benefits of the preferred option were quantified. The presentation of the estimates has been improved now, also to make clear that there is no double counting. In addition to the preferred option, the costs and benefits of the other options have been estimated as well. The distributional impact of the measures is explicitly included in the impact analysis; their effect on SMEs is analysed in section 8.1, and Annex 8 regarding the SME test is added. The burden reduction potential, for statisticians, respondents as well as data-holders, was already an element of the costs and benefits overview for the preferred option but it has been developed further for all options. These results of the impact analysis (documented in chapter 6 and Annex 3) have, of course, been used in the comparison of options and justification of

	the preferred option.
(8) The risks associated with the quality of privately owned digital data and skill shortages as well as the measures to mitigate those risks should be discussed in more detail.	In principle, the risks related to all costs and benefits are relevant to the impact analysis. For risks related to the quality of the reused data and skills required, the quality framework for European statistics applies, and the way to deal with these risks and the mitigating measures is discussed in chapter 8.

The technical comments directly received from the Regulatory Scrutiny Board have been addressed as well.

The Regulatory Scrutiny Board reviewed the revised impact assessment and gave a positive opinion with reservations on 27 March 2023. The impact assessment was revised as follows.

Comments of the RSB	*How and where comments have been addressed*
(B) Summary of findings	
(1) The report does not present the options in a way that brings out clearly the key policy choices.	The key policy choices have been clarified by an improved presentation and assessment of the related measures of the options.
(2) The report is not clear on what type of assessment will be undertaken to justify the inclusion of certain data collections from private actors.	A description is added of the process and assessment to be applied for adding statistics that are partly or completely based on new data sources.
(3) The mechanism to trigger the crisis-response measures is not sufficiently explained.	The explanation of the mechanism to trigger the crisis response measures has been elaborated.
(4) Some key assumptions used in the cost benefit analysis are not explained.	All key assumptions in the cost benefit analysis have been identified and discussed.
(5) The choice of the preferred option is not sufficiently justified to address effectively and efficiently each specific objective.	This point has been addressed together with (1) by a more elaborated assessment of the related measures of the options, in terms of each specific objective.
(C) What to improve	
(1) The presentation of measures proposed under each option should be improved to increase clarity; the measures which appear to be common for all or two	Measures that figured in more than one option have been disentangled. Measures of different options that are related are now presented in a way that clearly shows their

options should be presented in a coherent way. In addition, the rationale for having common measures should be reconsidered in several cases where there seems to be inconsistency with the proposed approach. The report should consider alternative combinations of measures to bring out clearly the available policy choices or explain why these are not relevant or clearly less performing than the two options presented. For instance, it should consider combining some policy measures of policy options 1 and 2, including for the specific objective 2 to react faster in time of crisis.	relationship (chapter 5). In some cases this had led to a reassessment of specific measures. For all related measures, the available policy choices are discussed and evaluated (chapters 6 and 7), thereby clarifying the key policy options and improving their consistency. The discussion now explicitly links the performance of the options and their measures to the specific objectives and shows that alternative combinations of measures would clearly result in inferior performance. This includes the policy measures of option 2 related to the responsiveness in times of crisis (see also (5) below).
(2) The report should better explain the process of including new data collections in the Annual Work Programme of the ESS, what type of assessment would have to be undertaken and whether this process would be different for the specific digital data collections from private data owners.	A description of the process of adding statistics to the AWP has been added, together with a specification of what needs to be done in addition in the case of statistics that are partly or completely based on new data sources (chapters 6 and 8).
(3) The report should explain how the crisis mode measures would be triggered. It should be clear under what circumstances, based on which criteria and under which decision making process the crisis mode is reached. Despite assuring that the initiative will be complementary to other crisis response legislation (e.g. the Single Market Emergency Instrument (SMEI)), it is not clear under what circumstances the ESS would respond to urgent data demands in times of crises. The report should explain which of the modes envisaged in the SMEI Regulation (if any) would trigger the application of the crisis response measures within the ESS.	It has been clarified that the aim is for the ESS to be responsive to information needs arising from a crisis situation that has been officially declared by an Institution but not to decide on its own on the existence of a crisis. It is only in cases where an emergency mechanism has been formally triggered in accordance with procedures established by Union law that Eurostat should have the capacity to organise in parallel a response at ESS level to meet urgent information needs arising from that crisis and when these needs have not already been addressed through the relevant emergency mechanism. It has also been clarified that in the case where the single market has been impacted by specific disruptions and shortages or possible intra-EU restrictions to the free movement of goods, services and persons, the ESS statistical response will only be initiated when the Single Market emergency mode has been activated by the Council and with the objective to meet information needs that cannot be covered by information requests addressed to representative

	organisations or economic operators in crisis-relevant supply chains (chapter 8).
(4) The report should explain and justify the assumptions used in the cost benefit analysis. It should explain how the numbers of expected crisis and ESS data sharing cases as well as new statistical domains were estimated. It should provide the justification for the different numbers of cases expected under each option. As those assumptions significantly impact the cost and benefit analysis, the report should undertake a sensitivity analysis and be clear about the level of uncertainty in the analysis.	The costs and benefits critically depend on assumptions on the numbers for each type of use cases. These numbers have now been explained and justified, taking into account that the revision concerns a framework regulation. The consequences of the intrinsic uncertainty of the assumptions have been assessed, which shows that the conclusions of the impact assessment remain valid as long as the number of use cases are within a very reasonable range of values around the assumed numbers (chapters 6 and 7).
(5) The results of the cost benefit analysis should be more transparently reflected in the justification of the preferred option. The report should explain why the policy measure 3.7 is not included in the preferred option package instead of the policy measure 2.7, as the report concludes that the policy option 2 is more effective and efficient then policy option 1 regarding the achievement of the specific objective 2 to provide mechanism and tools to react faster in times of crisis. The report should be more explicit that the preferred option is the most costly for businesses as regards measures related to crisis response and assess the corresponding impacts on competiveness. It should differentiate more clearly the technical feasibility of options from the support these received by stakeholders.	The disentanglement of the measures as described under (1) led to some adjustments of the measures. (chapter 5). Their impact has been assessed for each option (chapter 6) in a way that allows for the transparent comparison of the options (chapter 7). The criteria used for assessing the realisation of the specific objectives have been further clarified, in particular by distinguishing between technical feasibility and the estimated effect of political aspects (such as differences between measures in respect of subsidiarity) on feasibility. This made it possible to better explain why measure 2.7 rather than 3.7 is included in the preferred option,.

4. EVIDENCE, SOURCES AND QUALITY

Parts of the evidence and sources are described in other annexes. Annex 2 describes the consultations that have taken place, in particular the public consultation, the in-depth stakeholder interviews, the online survey, the online stakeholder validation workshop, and other targeted consultations. Annex 3 includes a summary of costs and benefits, which is partly based on evidence already gathered during the preparation of the impact assessment of the Data Act, or related to other statistical initiatives such as the intensification of data sharing in the context of making statistics on intra EU trade. The evidence, sources and quality of the summary of costs and benefits are described in Annex 3 as well.

More evidence has been gathered, within the ESS and the wider statistical community, and by using external expertise. This evidence, and its sources and quality, is described below.

Evidence gathered within the ESS and the wider statistical community

The decision by the Commission to start an initiative to revise Regulation (EC) No 223/2009 was taken after a long period in which the ESS gathered evidence of its necessity, especially in respect of its first specific objective, to exploit the full potential of digital data sources for official statistics. First of all, quite some research was done to explore the potential for official statistics of making use of data held by the private sector. A number of potential digital data sources were investigated and research carried out resulting in experimental statistics. One of the biggest such efforts was the so-called ESSnet Big Data, which was carried out from 2016 to 2021 on the basis of a multi-beneficiary grant agreement involving 28 partners, including the NSIs of 23 countries. The results[44] showed that the potential in terms of faster, more detailed and better statistics was very high indeed, but that the basis of voluntary partnerships between businesses and NSIs was by far insufficient to systematically sustain official statistics. Moreover, the potential benefits would not only comprise the statistical information produced, but also efficiency and a reduction of the need for statistical surveys. Lack of data access appeared to be one of the main bottlenecks to exploit relevant data sources and special attention was paid to the issue of data access[45]. The ESSnet Big Data was overseen by an external review board, to ensure the quality of the results.

From the research outcomes a number of use cases of potential use of new data sources for official statistics were derived; some of them were used for the impact assessment for the Data Act (see Annex 3). For the impact assessment for the revision of Regulation (EC) No 223/2009 these and other use cases were updated and used in the in-depth interviews with stakeholders (see Annex 2), to get a more concrete view on the obstacles, the potential costs to data holders and producers of statistics, and the potential benefits for, among other things, policymaking.

Within the ESS, a group of Directors-General of NSIs also investigated in 2020 the situation at national level in respect of legal access to privately held data. A questionnaire was designed with questions on a large number of aspects, from actual use of new data sources to obstacles to data reuse, from existing legal provisions to intentions to initiate legal changes, from data access during the pandemic to policy needs. Answers were obtained from 26 ESS Member States, on the condition of confidentiality. The results not only showed the limits of an approach based on voluntary partnerships, but also demonstrated the wide divergence at national level in data access practices and associated legal conditions. At this stage the need for a legislative initiative at EU level was evident to all NSIs, and in June 2021 the ESS Committee (the highest authority within the ESS) unanimously adopted a position paper on the future Data Act proposal about the need for access to privately held data for official statistics[46].

---

[44] https://ec.europa.eu/eurostat/cros/content/essnet-big-data-1_en

[45] For instance by organising a seminar with data holders on this subject, see https://ec.europa.eu/eurostat/cros/content/WP5_Meeting_2016_09_22-23_Luxembourg_Workshop_en

[46] ESS Committee position paper on the future Data Act proposal, June 2021.

Evidence was also needed concerning the choice of safeguards and conditions that should accompany the reuse of privately held data for official statistics. Since data access is crucial to statistical systems worldwide, data access principles received attention by the international statistical community at an early stage. The UN Global Working Group on the Use of Big Data for Official Statistics drafted a set of eight data access principles already in 2015, which were well received when presented at a conference with business sector participation[47]. Data access principles were also articulated by other organisations to which data sharing is relevant, such as the OECD[48], and the Commission has also provided guidance on such data sharing[49].

Use of external expertise

In April 2021, the Commission created the High-Level Expert Group on facilitating the use of new data sources for official statistics[50]. Its task was to provide recommendations aimed at enhancing data sharing between businesses and government (B2G) for the purpose of producing statistics (B2G4S), after a similar Expert Group provided recommendations on enhancing data sharing between the private and the public sector in general[51]. The B2G4S Expert Group consisted of 20 independent experts with various backgrounds that were particularly relevant to B2G4S, including from the business sector, research and academia, public administration and the civil sector. In June 2022 it produced its final report[52], entitled *Empowering society by reusing privately held data for official statistics, a European approach*. The report comprised an articulated set of recommendations aimed at enabling the sustainable reuse of data collected by the private sector for the development, production and dissemination of official statistics, including recommendations on safeguards and conditions [EG B2G4S p48-56]. This report is one of the main sources of evidence of this impact assessment.

For this impact assessment, the Commission contracted ICF SA, Belgium, to carry out a support study[53]. Apart from desktop research, ICF carried out the in-depth stakeholder consultation, the online survey and the online stakeholder validation workshop (see Annex 2). Thus, much of the effort was aimed at collecting evidence from stakeholders, not only on their views and positions, but also on the problems and their drivers, on the merits of the policy options and measures, and on experienced and expected costs and benefits. The wide variety and number of stakeholders consulted make the results reasonably robust, but the data that could be obtained on costs was limited, and should be considered anecdotal rather than representative. Therefore, the costs and benefits mentioned in Annex 3 are largely based on other sources (specified in that annex). Many

---

[47] Global Conference on Big Data for Official Statistics, Abu Dhabi, 20-22 October 2015.

[48] OECD (2021), Recommendation of the Council on Enhancing Access to and Sharing of Data, OECD/LEGAL/0643, October 2021.

[49] European Commission (2018), Staff Working Document, Guidance on Sharing Private Sector Data in the European Data Economy, 25 April 2018.

[50] https://ec.europa.eu/eurostat/cros/content/expert-group-facilitating-use-new-data-sources-official-statistics_en

[51] https://www.euractiv.com/wp-content/uploads/sites/2/2020/02/B2GDataSharingExpertGroupReport-1.pdf

[52] https://ec.europa.eu/eurostat/documents/7870049/14803739/KS-FT-22-004-EN-N.pdf/052b4357-bf8e-9ce4-c063-7e806c045dac?t=1656335798606

[53] ICF, *Study to support an impact assessment for the revision of Regulation (EC) No 223/2009 on European statistics, Final Report,* 18 November 2022.

parts of this impact assessment are partly based on evidence provided by the support study of ICF, in particular the scoring and comparison of the policy options. Eurostat has monitored the work of ICF on a weekly basis and considers the results to be of professional quality.

ANNEX 2: STAKEHOLDER CONSULTATION (SYNOPSIS REPORT)

1. Introduction

In the context of the impact assessment on the revision of Regulation (EC) No 223/2009 on European statistics, various consultation activities were conducted between February 2022 and November 2022. The purpose of the consultation was to collect evidence and views from a broad range of stakeholders, giving them an opportunity to provide relevant data and information on the problems and potential solutions concerning the challenges that European statistics are facing. While attempting to reach the widest possible range of stakeholders, the results of the consultation activities are not designed to be representative. This annex presents the results of the consultation activities carried out.

The consultation activities included:

- a call for evidence published on the "Have your say" portal and open from 21 February to 21 March 2022,
- an online public consultation conducted via a questionnaire published on the same portal and open from 19 July to 25 October 2022,
- stakeholder interviews carried out between October and November 2022,
- an online survey launched on 5 October and closed on 7 November 2022,
- an online stakeholder workshop carried out on 8 November 2022.
- High-level meeting of DGs of NSIs.

Concerning the call for evidence and the public consultation, those were open to the public.

2. Key stakeholders

For all stakeholder activities, the main stakeholder groups identified were:

1. Producers of European statistics
2. Users of European statistics, including
    - Institutional users of European statistics
    - Business users of European statistics
    - The media
    - The general public
3. Providers of primary data for production of European statistics, including
    - Public administrations
    - Individual data providers (respondents)
    - Businesses and other organisations
4. Other stakeholders, including

- Subjects to which data collected for the production of European statistics pertain

For **the call for evidence**, 21 answers were received coming mainly from EU citizens (8 respondents), followed by public authorities (4) and NGOs (4). Respondents to the call for evidence were mainly located in Netherlands (7), Germany (4) and Belgium (4).

For **the public consultation**, 204 answers were received coming mainly from EU citizens (83 respondents), followed by public authorities (71), others (15), academic/research institutions (14) and business associations (10). Respondents to the public consultation were mainly located in Spain (31), Germany (26) and Greece (25). The respondents were from 33 countries (26 EU Member States, Bangladesh, China, Iceland, Norway, Russia, Switzerland and United Kingdom).

For **the stakeholder interviews**, 30 in-depth interviews were conducted, 16 with institutional statistics users, 10 with data producers (NSIs), 3 with private data holders and 1 with an individual expert. The participants were located in Belgium, Estonia, France, Germany, Greece, Ireland, Italy, Netherlands, Norway, Poland, Spain, United Kingdom.

For **the online survey**, it received a total of 27 replies, 18 from NSIs and 9 from private data holders. The respondents were located in Austria, Belgium, Bulgaria, Cyprus, Czech Republic, Denmark, El Salvador, Finland, France, Greece, Italy, Hungary, Latvia, Lithuania, Netherlands, Norway, Poland, Slovenia, Spain, United States.

For the **online stakeholder validation workshop**, it gathered 26 participants representing 21 organisations – 5 NSIs, 8 statistical users, 4 business associations, 2 Private data owners, 2 NGOs. The respondents were located Belgium, Estonia, Finland, France, Germany, Italy, Netherlands, Norway, Poland, Portugal, United Kingdom, the United States.

Table 6: Stakeholder participation

Stakeholder type	**Public consultation and Call for Evidence on "Have your say"**	**Targeted consultation(in-depth performed by ICF)**	**Targeted consultation (online survey performed by ICF)**	**Final workshop**
Producers of European statistics	x	x	x	x
Institutional users	x		x	x
Business users	x		x	x
General public	x			

Public administrations	x	x	x	x
Individual data providers (respondents)	x			x
Businesses and other organisations (private data holders)	x	x	x	x
Businesses and other organisations (other)	x			x
Subjects of data collection	x			x

3. Summary of results

3.1 Call for evidence

Public authorities provided detailed feedback on the initiative. They mainly agree with the problems that need to be tackled, as identified in the Call for Evidence, and most dominantly refer to the need for access to new data sources. In that respect they mostly point to the need of ensuring that such access to privately held data should be free of charge for official statistics and that the inclusion of new data sources must not entail that the reliability and credibility of official statistics are compromised. Several public authorities also refer to the need of a strengthened coordination and possibility of sharing data.

A few EU citizens express concerns about citizens privacy, but one also points to it being essential for the compilation of statistics that there is access to relevant data sources, and that data holders should not be able to maximise their profit from this need. A few business associations point to the value of European statistics but point to the need of treating data with confidentiality when analysing it.

3.2 Online public consultation

The questionnaire of the online public consultation gathered feedback on the different measures considered in preparing the revision of the legal framework for European statistics. It notably touched on the following questions:

- What needs to be done to make European statistics fit for the future and more relevant to user needs?
- What are the most important factors in ensuring that European statistics better meet user needs?
- How important is it to make digital data held by the private sector available for the production of European statistics?
- What are the thematic areas of European statistics that might benefit the most from making data held by the private sector available?

- How important are a number of conditions for ensuring responsible use of digital data in European statistics (including conditions relating to only collecting for statistical purposes and to the minimisation principle)?
- Are European statistics sufficiently responsive to emerging user demand, including during public emergencies and crises?
- How can European statistics be made more responsive?
- For what purposes would easier and more systematic data sharing between statistics authorities be helpful within the ESS?
- What kind of conditions and safeguards should apply when sharing data within the ESS?

Of all respondents 3% (or 6[54] out of 204 respondents) replied that they considered their company/business a data holder, i.e., a company/business holding personal data or non-personal data that could be used or is known to be used for the production of official statistics. To the question, what needs to be done to make European statistics fit for the future and more relevant to user needs, respondents could select one or more options. 70% of respondents (or 143 out of 204 respondents) considered it most important to combine sources to provide more and better insights into economic and societal developments, 66 % (or 135 out of 204 respondents) considered it most important to provide more granular statistics (e.g. for social groups and territorial units), and equally 66% (or 135 out of 204 respondents) of respondents considered it most important to provide more up-to-date statistics, e.g. through flash estimates and more frequent statistics. According to 57% of respondents (or 117 out of 204 respondents) it was most important to respond faster to emerging demand for data and statistics, especially during public emergencies and crises, to 56% (or 114 out of 204 respondents) it was most important to improve the way statistics are disseminated and communicated, to 51% (or 105 out of 204 respondents) it was most important to provide more statistics about new phenomena, equally to 51% (or 104 out of 204 respondents) it was most important to provide more metadata, explaining the statistics, keep a change log and notify the most recent update, and finally to 32% (or 66 out of 204 respondents) it was most important to publish more statistics under development, to engage more intensively with users of European statistics.

To the question, what are the most important factors in ensuring that European statistics better meet user needs, respondents could select one or more options, where the most important options were considered to be as follows:

- Sustainable access to relevant data sources to be used for production of European statistics (75%) (or 154 out of 204 respondents); broken down by selected stakeholders, this was considered important by 56 out of 83 EU citizens, 61 out of 71 public authorities, and 2 out of 5 businesses
- Modern IT infrastructure to support the production and dissemination of European statistics (65%) (or 133 out of 204 respondents)
- Sufficient resources for the ESS (60%) (or 123 out of 204 respondents)
- Reskilling and upskilling of staff in national statistical authorities and Eurostat (50%) (102 out of 204 respondents)

[54] However, only 15 respondents out of 204 identified themselves as either “company/business organization” or “business association”

- Better statistical processes (34%) (or 69 out of 204 respondents)

On the question of the importance of making digital data held by the private sector available for the production of European statistics, 83% (or 169 out of 204 respondents) of the respondents considered this being of very high or high importance. Broken down by selected stakeholders, this view was supported by 69 out of 83 EU citizens, 64 out of 71 public authorities, and 3 out of 5 businesses.

According to respondents the thematic areas of European statistics that might benefit the most from making data held by the private sector available, are the area of Economy and Finance (70%) (or 142 out of 204 respondents) followed by the area of Environment and Energy (60%) (or 123 out of 204 respondents).

To the question of the importance of different conditions for ensuring the responsible use of digital data in European statistics, respondents identified the following factors:

- The statistics authority and the data holder should ensure, where relevant, data privacy and confidentiality (99% of very high or high importance) (or 201 out of 204 respondents)
- The statistics authority and the private data holder should practise full transparency towards the public and the people the data relate to (87% of very high or high importance) (or 177 out of 204 respondents)
- The statistics authority and the data holder should be mutually obliged to collaborate in good faith (86% of very high or high importance) (or 175 out of 204 respondents)
- The request for access to data should explain why the data are necessary or useful for compilation of official statistics (79% of very high or high importance) (or 162 out of 204 respondents)
- Mechanisms should be in place to address potential disagreements over data requests between the data holder and the statistics authority (77% of very high or high importance) (or 157 out of 204 respondents)
- The reputation and business interests of the data holder should be respected and safeguarded (77% of very high or high importance) (or 156 out of 204 respondents)
- Data made available for the production of official statistics should be used only for that purpose (68% of very high or high importance) (or 139 out of 204 respondents)
- The statistics authorities should request only the minimum data they need (minimisation principle) (58% of very high or high importance) (or 119 out of 204 respondents)

11% of respondents (or 23 out of 204 respondents) consider that European statistics are sufficiently responsive to emerging user demand, including during public emergencies and crises, whereas 72% (or 146 out of 204 respondents) consider that European statistics are somewhat responsive, but not enough, and 8% (or 16 out of 204 respondents) consider those statistics not responsive. When considering the share of selected stakeholders supporting the dominant view, that European statistics are somewhat responsive, but not enough, this was supported by 53 out of 83 EU citizens, 58 out of 71 public authorities, and 3 out of 5 businesses.

According to 56% of respondents (or 115 out of 204 respondents) European statistics could be made more responsive by means of a more intensive use of digital data to follow fast societal, environmental and economic changes, whereas 50% of respondents (or 103 out of 204 respondents) considered it could be achieved by making more and better use of already existing statistical data to respond to demand. 49% of respondents (or 99 out of 204 respondents) considered that it could be achieved through more collaboration with research and the academic sector, 47% (or 96 out of 204 respondents) considered that it could be done through more coordination at European level to better react to crises and emergencies, 38% (or 77 out of 204 respondents) considered that it could be achieved by more experimentation and use of statistics under development for greater engagement with users, 31% (or 63 out of 204 respondents) considered that it could be achieved through more dialogue with users, and 30% of respondents (or 62 out of 204 respondents) considered it could be achieved through more resources invested in dialogue, collaboration and experimentation.

As regards purposes for which easier and more systematic data sharing between statistics authorities would be helpful within the ESS, 72% of respondents (or 147 out of 204 respondents) consider it helpful to reduce response burden and allow for reuse of already collected data, 69% (or 140 out of 204 respondents) consider it helpful to increase the quality of official statistics, 65% (or 132 out of 204 respondents) consider it helpful to garner synergies and cost efficiency in production of official statistics at EU and national levels, 63% (or 128 out of 204 respondents) consider it helpful to enable production of cross-border official statistics that cannot be compiled correctly as a sum of national estimates, 61% (or 125 out of 204 respondents) consider it helpful to increase potential for research in official statistics (i.e., developing new methods for compiling official statistics), and 56% of respondents (or 115 out of 204 respondents) consider it helpful to help develop new statistics (including for cross-border regions).

To the question, what kind of conditions and safeguards should apply when sharing data within the ESS, 75% of respondents (or 154 out of 204 respondents) consider that purpose limitation for sharing personal data (sharing of data only for agreed purposes) should apply, 72% (or out of 204 respondents) consider that effective protection of data using state-of-the-art technologies, should apply, 45% (or 92 out of 204 respondents) consider that technologies to minimise the need to transmit data should be used, 31% (or 63 out of 204 respondents) considers that purpose limitation for sharing non-personal data (sharing of data only for agreed purposes) should apply, and 8% (or 16 out of 204 respondents) consider that no conditions should apply.

Finally, on the extent to which respondents agreed to a number of statements, the results were as follows:

- Statistics authorities should set standards for interoperability (86% of respondents agree or strongly agree) (or 176 out of 204 respondents)
- Statistics authorities should develop and maintain a catalogue of data assets in their national data ecosystem, and make the catalogue publicly available (86% of respondents agree or strongly agree) (or 176 out of 204 respondents)
- Statistics authorities should provide professional advice to organisations within their ecosystem on issues related to data and data processing, such as quality, data reuse, intellectual property, confidentiality, security and

metadata (85% of respondents agree or strongly agree) (or 173 out of 204 respondents)

- Statistics authorities should assess the quality of statistics made available to the public by other organisations (75% of respondents agree or strongly agree) (or 152 out of 204 respondents)
- Statistics authorities should mediate among organisations interested in data sharing and reuse (61% of respondents agree or strongly agree) (or 124 out of 204 respondents)

3.3 Targeted consultations

In-depth stakeholder interviews

ICF, preparing the study in support of this impact assessment, conducted 30 interviews between October and November 2022 primarily to inform use cases and to understand statistical users' data needs and the potential benefits associated with enhanced B2G4S data sharing (see chapter 6 and 7 of the impact assessment). The contractor interviewed 10 data producers (NSIs), 3 private data holders, 16 statistical users' organizations and 1 individual expert [ICF, Annexes 2 and 3].

Online survey

The online survey, equally conducted by ICF, targeted NSIs and private data holders. NSIs were reached by Eurostat through the Directors of Methodology (DIME/ITDG) network. ICF reached out to 69 individual companies and 81 business associations with a focus on European affairs (who were asked to share the survey with their members). The survey received a total of 27 replies, 18 from NSIs and 9 from private data holders.

Integrated results of targeted consultations

The direct or indirect benefits as identified by various stakeholders in various use cases are summarised in Table 7:

Table 7: Results of targeted consultations

Use case	Stakeholder type	Direct/indirect benefits
Mobile network operators	Society	Less overall burden on businesses as automated data sharing mechanisms replaces lengthy surveys
	Statistical users	Fine-tuning facilities and services; better decision-making including better understanding of interaction between territories, human behaviour, spatial planning
	NSIs	Better data in terms of quantity, quality, granularity, and timeliness; Enhanced reputation
	Private data holders	Compensation: a three-year contract (total value EUR 300.000), subsequent annual contracts (EUR 24.000); Involvement in research and scientific publications; Sharing of methodological expertise; Increased data quality might lead to monetisation and commercialisation of data; Enhanced reputation;
Financial transaction data	Society	Reduce response burden for the business sector by saving up to 3.400 hours of work

Use case	Stakeholder type	Direct/indirect benefits
	Statistical users	Production of short-term statistics about the retail trade sector and certain B2C service industries Larger coverage and bigger transaction datasets
	NSIs	Cut up 20,000 individual questionnaires per year; Higher data timeliness and frequency
	Private data holders	(Vipps) Compliance with national Statistics Act (Fable) Supporting the company's mission Bringing some "gravitas" to their datasets as a result of collaboration with respected public institutions
Smart meters	Statistical users	Better understand changing patterns of energy demand and consumption; To produce better overviews of specific markets and fill knowledge gaps; Better city planning, including residence and mobility
	NSIs	Faster access to more accurate data; Smart meter data go far beyond what would be possible to achieve through traditional data gathering methods
	Private data holders	-
Smart devices (IoT)	Statistical users	To understand levels of pollution within and across regions; Identify and improve issues within a large city
	NSIs	Frequent and more detailed data, improving accuracy; innovation
	Private data holders	-
Collaborative economic platforms (tourism)	Statistical users	Filled information gap related to knowledge of short-term rentals, allowing a more accurate debate about tourism; Better understanding on how travelling is developing and how is impacting tourism
	NSIs	-
	Private data holders	-

Equally the following Table 9 gives a summarized overview of various stakeholders estimate of costs in various use cases:

Table 8: Stakeholders' estimate of costs in various use cases

Use case	Stakeholder type	One-off/recurrent costs	Comments
Mobile network operators	Statistical users	-	(If data sharing **did not** occur): information gaps leading to poorer decision-making; additional time to find alternative resources; additional procurement to access commercial databases
	NSIs	Time/human resources for data processes and analysis (one-off) Compensation to MNOs (one-off/recurrent)	NSIs need appropriate resources to handle data coming from MNOs due to significant pre-processing and cleaning before data is usable to generate official statistical purposes
	Private data holders	"Minimal" human resources for data processes and analysis (one-off)	

Use case	Stakeholder type	One-off/recurrent costs	Comments
Financial transaction costs	Statistical users	-	(If data sharing did not occur): information gaps leading to poorer decision-making; additional time to find alternative resources; additional procurement to access commercial databases
	NSIs	"Very low" cost in setting up data sharing (one-off); Time dedicated to establish dialogue with company and explain the legal basis: 2/3 months full time work by a team of 5/6 individual (one-off)	Data sharing was pre-existing so previous technological infrastructure was in place
	Private data holders	(Vipps) Development costs: one full week (one-off) Daily data/report delivery: 1-2 months full time work (recurrent) (Fable) Long time to set up the collaboration (one-off) "Minimal" technological infrastructure (i.e. server) costs (recurrent)	(Vipps) Data sharing was pre-existing so previous infrastructure was in place (Fable) Costs might increase if data sharing grows in complexity, involving third parties, etc.
Smart meters	Statistical users	-	(If data sharing did not occur): information gaps leading to poorer decision-making; additional time to find alternative resources; additional procurement to access commercial databases
	NSIs	"Low" personnel costs, significantly "minimal" in comparison to collecting such data in a traditional way (recurrent) Hiring new data scientist (recurrent) Low initial costs setting up a server (one-off)	
	Private data holders	-	
Smart devices (IoT)	Statistical users	-	(If data sharing did not occur): information gaps leading to poorer decision-making; additional time to find alternative resources; additional procurement to access commercial databases
	NSIs	"Low" costs associated with personnel time in establishing partnership (one-off). Team was later expanded into small number of full-time roles (recurrent).	NSI noted that securing certain types of data from private owners can be "prohibitively expensive"
	Private data holders	-	-
Collaborative economic platforms (tourism)	Statistical users	-	(If data sharing did not occur): information gaps leading to poorer decision-making; additional time to find alternative resources; additional procurement to access commercial databases
	NSIs	-	NSI commented that NSIs should not pay to reuse privately held data
	Private data holders	"Manageable costs" (if there is one data portal, one API, and one format) (recurrent)	Costs become "unmanageable" if the company had to share the data with each Member State or share it at very short intervals

3.4 Online stakeholder workshop

The objective of the online stakeholder workshop, which took place on 8 November 2022, was to seek stakeholders' validation of the findings of the preceding consultation activities. The invitees were therefore overlapping with the stakeholders having been subject of the targeted stakeholder interviews and the online survey. The workshop included an audience interaction using Mural where the participants could express their support for various policy options and measures and discuss their potential impacts, including on SMEs and on public authorities and their budgets.

When looking at the potential effectiveness of the envisaged policy options, stakeholders participating in the validation workshop suggested that policy option 1 of the impact assessment would have the highest level of effectiveness to achieve the objectives of the revision (56% voted 'High' and 28% 'Very high'). When looking at perceived efficiency and perceived coherence, it was also policy option 1 that was considered as having the highest level of, respectively, efficiency to achieve the objectives of the revision (50% voted 'High' and 11% 'Very high') and coherence in relation to other EU initiatives/regulation in this field (50% voted 'High' and 22% 'Very high'). However, when looking at the perceived EU added value of the envisaged policy options, stakeholders participating in the validation workshop suggested that policy option 2 would have the highest level of EU added value compared to Member States acting separately to achieve the same results (44% voted 'High' and 39% 'Very high'). Policy option 1 was also deemed as having a high level of EU added value as 44% of workshop participants voted 'High' and 33% voted 'Very high'.

3.5 Other targeted consultations

In the anticipation of the adoption of the Data Act proposal, the topic of access to privately held data for statistical purposes was discussed within the ESS Committee at several occasions and led, amongst others, to the adoption of a position paper on the importance of access to privately held data for official statistics in June 2021[55].

On 20 and 21 October 2022 the Conference of European Statistics Stakeholders took place at the University of Rome, "La Sapienza". This conference was organised by the European Statistical Advisory Committee (ESAC), the European Central Bank (ECB), the Federation of European National Statistical Societies (FENStatS), and Eurostat. The Expert Group on facilitating the use of new data sources for official statistics presented its recommendations, and Eurostat presented the initiative to revise Regulation (EC) No 223/2009. The recommendations of the Expert Group and the initiative to revise Regulation (EC) No 223/2009 were very well received.

3.6. Under the **French Presidency a high-level meeting was organised in Lyon on 7-8 April 2022** on 'Making the European Statistical System fit for the future'. Subsequently a dedicated meeting of the Presidents and Directors-General of the NSIs was held in Luxembourg on 18 and 19 May 2022 to discuss the conclusions.

---

[55] Cf. "European Statistical System (ESS) position paper on the future Data Act proposal ¬ Access to privately held data is urgently needed for producing new, faster, more detailed official statistics", June 2021.

At the meeting in Lyon, a large consensus emerged on a number of topics to be covered by the revision of the Regulation, including the provision of a legal basis for sustainable access to privately held data for official statistics, the use of new technologies and the development of experimental statistics, the recognition of the possibility for NSIs to assume new roles and tasks, and the fostering of data sharing in the ESS.

The objective at the subsequent meeting of the ESS Committee in Luxembourg on 18 and 19 May 2022 was to identify more precisely how the proposed changes could be addressed at operational level and be effectively implemented for two issues, namely sustainable access to privately held data and data sharing in the ESS for producing European statistics.

During the discussions, the urgent need to set up a proportionate, limited, predictable framework for making data available for the compilation of official statistics was highlighted: the objective is to set up a mechanism that is solid from a legal point of view and that will effectively oblige data holders to make their data available to NSIs and Eurostat, under appropriate safeguards and conditions. The importance to address privately held data as another data source was also emphasized. At the same time, it was recognised that not all details of data sharing arrangements should be set in the law.

It was also considered that data sharing within the ESS should be strengthened, notably in relation to cross-border phenomena, even though different views were expressed on how to best achieve this objective with some participants noting that current Regulation (EC) No 223/2009 already allows for data sharing on a voluntary basis and the possibility to address this issue in sectoral legislation was mentioned as an alternative. Views however broadly converged on the need to ensure that any strengthening of data sharing should be limited and purpose oriented, based on high security controls within a delineated common data space and built on modern data access technologies such as privacy enhancing technologies.

ANNEX 3: WHO IS AFFECTED AND HOW?

1. PRACTICAL IMPLICATIONS OF THE INITIATIVE

Depending on the specific objectives and related measures, different types of stakeholders would be affected by the initiative.

Specific stakeholders can have various roles, e.g., businesses can be respondents providing data to the statistical system but can also be users of official statistics products. The following list stakeholders therefore elaborates on their (possible) role as this determines the impact of the measures on them.

As main stakeholders, we distinguish between the

- **The European Statistical System (ESS)**,
  i.e., Eurostat, the national statistical authorities (including the NSIs and the other national authorities) that produce European statistics.

- **The European System of Central Banks**, whose members produce monetary, banking and finance European statistics.
- **The European businesses and their organisations** distinguished according to their size into SMEs and large enterprises;
- The key institutional users at EU, national and regional level that make informed decisions and policies
- **The academic sector** and researchers
- **Media**, which take the role of informing the society and different parts thereof;
- European citizens and the **society** at large.

The European Statistical system (ESS)

The problem identified and measures to address them have a direct impact on ther daily work of the producers of European statistics. They collect data from different relevant data sources, process the data and produce European and national official statistics, which they disseminate to the public at large. In addition, they further develop European statistics and under strict conditions provide access to detailed but anonymized statistical data for research purposes. In the context of the revision of the Regulation on European statistics, Eurostat will be enabled to increase substantially its capability to provide insights and actively supports the political decision-making at EU level, sustains and fosters policymaking in the context of the EU political agenda, facilitates the implementation of EU policies and legislation, and promotes transparency and democratic accountability on EU and national policymaking to the public at large.

The European System of Central Banks

National central banks and the European Central Bank are part of the European System of Central Banks. They have similar tasks and roles as statistical authorities within their field of competence. They act as producer of European statistics but do also use European statistics for economic and financial analysis to base financial measures on sound data. They will benefit from increased richness and quality of European statistics. The central banks will substantially profit from higher agility of the statistical system in times of crises.

European businesses and their organisations

European businesses have a double role in official statistics. On the one hand, they participate in the data collection process as respondents. On the other hand, they are using official statistics as part of their decision-making process. As respondents, they are affected by any increase or decrease of response burden. It is in their interest to minimize and automate data requests. As users of official statistics, they will profit from receiving more timely and detailed data. Use of new data sources will affect a small portion of businesses disposing of large data assets, such as MNOs, banks, energy providers, or internet platforms. They might incur additional costs or might have certain additional obligations in order to enable statistical authorities accessing their data. The very large majority of businesses will rather use improved statistics based on these new data sources. They will benefit from faster data availability in cases of crisis. Multinational businesses holding large datasets will profit from burden reduction via a one-stop shop, i.e., one statistical authority assessing data for the ESS. They will also benefit from data stewards in statistical authorities who are able to professionally manage data partnerships. The other businesses will benefit because new data sources have the

potential to diminish response burden through reducing reporting obligations. Data assets are rather held by large enterprises also considering the scaling effects of providing services via the internet. Data start-ups and data intermediaries might be of smaller size and might be affected disproportionally by requests for data access for the purposes of official statistics. However, these requests are rather unlikely due to the limited capacity of these enterprises for generating and managing large data assets that would qualify for reuse by official statistics. On the contrary, they might be partners in developing and implementing data analysis and processing routines in data reuse scenarios with compensation for their services.

Key institutional users

Key institutional users and Policy makers need high quality, i.e., timely, accurate, comparable, harmonised data to be able to design, execute, monitor and evaluate policies and to take informed decisions. They will benefit from increased offer of data through use of privately held data, intensified data sharing within the ESS, and higher agility in case of urgent demands in times of crises. Statistical authorities provide statistics in an impartial way, which enables policy makers to assess and balance burden and benefits in their policy decisions. Dedicated data stewards are able to professionalize and rationalize reuse of data in the public sector. In addition, staff in governments will benefit from higher data literacy through training and education programs by statistical authorities. On the other hand, they have an interest in reducing administrative burden on businesses, and in reducing government spending. This can be achieved through efficient data sharing mechanisms, adjusting demands for data to their intended use, intensifying data sharing within the ESS, etc.

Public administrations as data providers

Public administrations can act as data providers (of so-called administrative data) for statistical authorities and/or use official statistics for their purposes. In recent years, administrative data has become more important as source for official statistics with the potential of minimizing burden on businesses and citizens. This development will become even more important with increasing degree of digitalisation of the public sector. In case of reuse of administrative data, burden on public administrations might increase with increased use of their data for the purposes of official statistics. On the other hand, statistical authorities will contribute to the professional management of the data and increase quality of these data through data stewardship functions. Staff in public administrations might also benefit from training and education programs organised by statistical offices as part of their possible functions within the public sector. Like policy makers, public administrations will benefit from timelier, more frequent and more granular data.

Academic sector and researchers

Researchers are users of data disseminated by statistical authorities but also contribute to further develop methods for processing and analysing data, which are applied by statistical offices. Researchers will benefit from enhanced statistical information with improved quality.

Media

The media are intermediaries between organisations producing data and information and citizens consuming that information. The media will benefit from increased offer of timely, adequate, granular, accurate and harmonised statistical data, which is disseminated impartially. This aspect is of utmost importance for public policy debate and for fighting fake information. There is also a strong demand for more granular data to satisfy the demand of media to serve specific users. Developing new tools and methods for communication and dissemination of statistical data following the changing habits of citizens are crucial in this respect. Media will specifically benefit from mechanisms responding to urgent user demands, e.g., in cases of crisis.

Citizens and society at large

There is a risk that official statistics will lose relevance for citizens through focussing too much on averages, larger populations and on national or European levels and thus not reflecting the data to daily experiences and living conditions of citizens. The citizens will benefit from more timely and granular information offers through increased reuse or privately held and administrative data by statistical authorities together with tailored communication services that might be delivered through private companies. Their role of data users is contrasted with their role of providers of data, either in traditional ways through statistical surveys or as users of digital services producing data assets at the service provider. These data are often personal and very detailed. Statistical offices ensure anonymization and ethical use of these data by legislation and by application of ethical codes of practice. Through developing and integration of privacy enhancing and preserving technologies in the statistical production process, misuse of personal data can be ruled out on technical basis. Increased reuse of new data sources will diminish response burden on citizens while ensuring anonymity of the resulting statistics. However, increased data sensitivity calls for increased participation of citizens in the process of prioritisation of user demands for specific statistical products.

2. SUMMARY OF COSTS AND BENEFITS

The application of use cases for the calculation of costs and benefits

The summary of the costs and benefits is based on the analysis and the developed assessment methods described in Annex 4, analytical methods[56]. Costs and benefits incur with implementations of the additional policy measures (use cases) based on the amended framework regulation. These include cases of reuse of new data sources, which have emerged as by-products of digital services, of additional data collections based on the urgent user demand motivated by crises, and mandatory mechanisms for internal ESS data sharing. Some of the additional cases of reuse are assumed to be implemented at European level, in which Eurostat would access data of data holders which operate at European-wide level, centrally to produce European statistics.

In order to quantify the impact of the different policy options, we assumed for each option a specific number of use cases that would be realised within a time span of 10 years and that would improve existing statistics or would produce new statistical output. The quantification follows the specific measures that are proposed under each of the options. In the case of crisis response, we assume that measures would be of temporary

[56] Annex 4, page 81 contains further references to literature used in this context.

character, i.e. direct costs and benefits would only occur during the time the measures would be executed. Indirect savings and benefits, e.g. through better treatment of crises and containment of negative effects on society and economy would not be limited in time. The quantification depends on the assumed number and type of use cases for each option. Although the numbers that are assumed are based, among other things, on extensive experience with pilot projects and partnerships between NSIs and data holding businesses, the assumptions are essentially informed expert guesses that may be too high or too low. This is unavoidable, given the fact the Regulation (EC) No 223/2009 is a framework regulation and the actual decisions on use cases are taken in the context of the Annual Work Programme as described in chapter 8. While the assumptions have a degree of uncertainty, they are realistic, as it is not so much the realisation of the given numbers of use cases that is uncertain, but rather whether this takes more than 10 years or less. The description of the impact of the policy options and their comparison (chapters 6 and 7, respectively) takes this into account.

It is important to note that the ESS will have to invest in methodological and quality frameworks as well as in providing staff of statistical offices with the required knowledge and skills set to use new data sources together with traditional sources and integrate them into the production of European statistics. Those investments in the quality aspects and in skills reinforcements are needed in all three options.

Categories of costs and benefits[57]

Organisational costs include costs related to entering into agreements, the monitoring of their execution, as well as additional consultations regarding logistics associated with data transfers, support, adjudication mechanisms, and other matters.

Methodological development costs relate to familiarisation with the dataset and relevant metadata on an ongoing basis, and the development and maintenance of methodologies to facilitate and enable the transformation processes from source data to statistics.

Infrastructure development costs cover mainly computing resources, the implementation of IT solutions: network configuration, hardware and software tools necessary for the handling of the new data and the processes that utilise them, including the writing of code and carrying out computations in end-to-end tasks.

Operational costs are related to the implementation of the production process, to quality assurance and dissemination activities, as well as to post-production product awareness and user support.

Upfront costs relate to efforts entailing all kinds of frontloaded costs, including exploratory and investigative research, as well as resources for negotiations, subject matter, legal, and ethical considerations and preparations.

Compensation of the data source relate to the fact that shared data may come from a wide variety of sources and may be put to various uses. This can lead to compensation requests. Their exact nature or the negotiated terms cannot be known a priori. Therefore,

[57] See: Sciadas G and Stavropoulos P., Methodological support to impact assessment of using privately held data by official statistics, Literature review and model, December 2021

the model allows for extensions to incorporate compensation to the source as and if necessary.

Improved quality of the existing statistical outputs refer to improving quality parameters, such as relevance, comparability, accuracy or timeliness as an effect of exchanging data or adding additional data sources to compute statistical outputs.

Extending the existing line of output or producing **new statistical products** may be based on use of new data sources or as a result of additional data sharing within the ESS. Here, new data sources are likely to be used in a multi-purpose approach, e.g. electricity smart meter data may be used for producing statistics on energy consumption at very granular level but also contribute to tourism statistics (occupation of secondary homes) or be used in the context of population and housing statistics.

Savings refer to savings caused by reducing survey sizes due to including data from new data sources.

The transformation of the shared data to products by the statistical office will lead to the realisation of societal benefits, while costs are calculated assuming an end-to-end production process under the control of the statistical office. The data holders may also participate in the production process. To the extent that they assume an assigned role in the overall production due to pushing out computation for a variety of reasons, they may be compensated. This will have no effect on the computation of the total costs; it will just determine their fair apportionment among cooperating entities. There will be upfront costs, which will be amortized over time and therefore our conceptualization focuses on recurring costs. However, upfront costs are included in the model.

Costs occur at both sides, the statistical offices as well as on the side of the businesses. They are distributed on the basis of estimated ratios, e.g., preparation costs are distributed equally between businesses and statistical offices, while costs for methodological developments are mainly paid by statistical offices. Savings for statistical offices and effects of burden reduction due to smaller sample sizes are equally distributed between statistical offices and businesses. Compensations are costs for statistical offices and benefits for the businesses. Compensation for data holders should partially compensate for additional costs at the data holder for preparing data for re-use by statistical offices according to agreed standards.

Who will be affected by reuse of innovative data sources

The European Statistical System has performed a number of projects to prepare for reuse of innovative data sources. In the ESSnet Big Data I and II[58], which ran from 2018 – 2021, national statistical institutes developed methodological and quality frameworks, and prepared prototypes of reuse of innovative data sources for improving current or producing new statistics. As possible data sources, the projects identified smart meters, automatic identification systems data of vessels and airplanes, mobile network operators' data, earth observation data, financial transaction data, data from internet platforms, smart farming. In all for the mentioned examples, the number of enterprises holding significant amounts of data is quite small. For example, Eurostat has committed a study

[58] Cf. https://ec.europa.eu/eurostat/cros/content/essnet-big-data-i_en, https://ec.europa.eu/eurostat/cros/content/essnet-big-data-1_en

on relevant platforms offering jobs online. Across Europe, approximately 1000 potential websites were identified. Currently, data is retrieved from 600 websites within the European Economic Area and Switzerland. The number of websites indicates the maximum number of enterprises running those sites. For mobile network data, 3 to 4 communication service providers cover the national market within Europe[59], which makes up 81-108 enterprises across the European Union. According to Eurostat structural business statistics, 90 providers of wireless telecommunication services covered 92% of the total turnover of all enterprises within the European Union in 2021[60]. Following these typical examples, it can be assumed that less than 10,000 relatively large **enterprises** across Europe will be affected by mandatory data sharing requests as reflected in data from European business statistics.

The baseline option (policy option 0 - PO0)

Based on experiences within the ESS, we assume one case of intensified data sharing among the members of the ESS. The assumption is motivated by the introduction of Intrastat, the statistics in intra EU trade, which started in 2012 with a prototype project[61] and took almost a decade to be fully implemented in the European context, where a series of organisational, methodological, procedural and governance issues had to be solved, and appropriate solutions had to be developed and implemented at Member States as well as at European level.

In addition, we assume one case of reuse of privately held data realized at national levels and that 18 out of 27 Member States would be successful in reusing new data sources. High initial investments on the side of NSIs constitute an obstacle to B2G4S, which, however, can be lowered through coordination and support at ESS level. This assumption is based on ESS experience with reusing platform data in the context of the collaborative economy, where methodological constraints limit the use of the data for producing European statistics.

Finally, we assume one case of crisis response, in which Eurostat would directly collect data from enterprises to produce European statistics and distribute intermediate data to the members of the ESS (data hub function). Data production initiatives during the COVID-19 crisis have revealed the complexity of coordinating national inputs with different levels maturity of data to produce new statistics at required quality level as regards coverage, coherence or comparability.

The development of guidelines is financed as part of the expenditures on methodological developments of Eurostat. Incentives to increase data sharing are as well financed as part of the annual budget of Eurostat and the members of the ESS.

---

[59] Cf. GSM Association: Mobile market structure and performance in Europe, February 2020, https://www.gsma.com/publicpolicy/wp-content/uploads/2020/01/GSMA-Mobile-Market-Structure-and-Performance-in-Europe_February20.pdf

[60] Enterprise statistics by size class and NACE Rev.2 activity (from 2021 onwards) [SBS_SC_OVW__custom_5134279], extraction 17/12/2022, https://ec.europa.eu/eurostat/databrowser/bookmark/3a05894b-145c-4d6a-809a-73cd585e4c71?lang=en

[61] Cf. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Intra-EU_trade_-_exchange_of_micro-data#Towards_an_EU_regulation_and_national_implementations

Figure 3: PO0: Application of statistical products / use cases

	A	B	C	D	E	F	G
1	Direct cost and benefit in Million EUR	B2G4S		Crisis response		Total direct benefits and costs	
2		at national level		at EU level (Eurostat as hub)		Total	
3		ESS	Businesses	ESS	Businesses	ESS	Businesses
4		1 Stat. products / domains		1 Stat. products / domains			
5	Preparation costs	1.8	1.8	0.0	0.0	1.8	1.8
6	Organisational	2.2	0.5	0.0	0.0	2.2	0.6
7	Meth. Development	9.7	1.1	0.2	0.0	9.9	1.1
8	Infrastructure	9.5	4.1	0.2	0.1	9.7	4.1
9	Operational	7.6	3.2	0.2	0.1	7.7	3.3
10	**Total costs**	**30.7**	**10.7**	**0.7**	**0.2**	**31.4**	**10.9**
11	Quality improvements	12.6		0.1		12.7	
12	Savings /burden red.	31.5	31.5			31.5	31.5
13	Additional outputs	10.8		0.8		11.6	
14	Compensation		0.9		0.0		0.9
15	**Total benefits**	**54.9**	**32.4**	**0.9**	**0.0**	**55.8**	**32.4**
16	**Net benefits / costs**	**24.2**	**21.7**	**0.3**	**-0.2**	**24.5**	**21.5**

Table 9: Cost and benefit for PO0: Reuse of new data sources and crisis response

The overall balance for the two cases is positive with a net benefit of EUR 24.5 million (F16) for the statistical system and EUR 21.5 million (G16) for businesses. The benefits for businesses mainly result from burden reduction of the B2G4S case due to decreasing sample sizes affecting the entire business sector, while it is assumed that the crisis response mechanism would produce new statistics, which would not lead to burden reduction. For both cases, data holders would face additional burden, which is at least partially compensated by the public sector. The crisis response mechanism would create additional cost of EUR 0.2 million (E16) on data holders.

	A	B
1	ESS data sharing	
2	1 Stat. products / domains	Million EUR
3	Burden reduction	39
4	Savings through running a central system	3

5	Costs of running a central system	-0.7
6	**Total**	**41.3**

Table 10: PO0: Cost and benefit: Voluntary ESS data sharing

The voluntary data sharing for one statistical data collection would result in burden reduction of EUR 39 million (B3). The burden reduction effects are due to the reduced number of enterprises in the data collection and less information asked to enterprises. In the case of EU Intratrade statistics, most enterprises are only asked for their exports instead of exports and imports. The total savings are estimated at EUR 41.3 million (B6).

In conclusion, the baseline option would not realise the benefits through intensified data sharing, the mechanisms for responding to urgent user demands, and the reuse of privately held data to cope with the increasing demand for European statistics. Statistical offices would tend to intensify use of administrative data for statistical purposes and try to increase the number of surveys. As a likely effect, statistical offices would increasingly suffer from already declining response rates to surveys, which would have in return a strong negative impact on the quality of the produced statistics introducing bias and lessen the quality dimension 'accuracy'. Statistical offices would continue investing in automation of surveys, i.e., use tools for automated data collection from enterprises or smart devices for supporting social survey data collection. There would be a need for investments in the coming years for these purposes while the benefits would be considerably lower than for the preferred option.

These negative developments will likely be somewhat reduced because of the increase in the availability of open data and voluntary data sharing that may result from the implementation of the European Data Strategy. However, this strategy does not foresee and is not expected to result in the sustainable and harmonised reuse of privately held data for official statistics, which is needed for the cost-benefit ratio to improve and to contribute to meeting user needs.

Policy option 1 (PO1)

For policy option 1, a total of 20 cases of reuse of privately held data (B2G4S) and cases of crisis response would be realised within the next 10 years. Out of these 20 cases, 15 would be B2G4S use cases and 5 would be cases of urgent user demands in times of crises. Out of the total of these 20 cases, 5 would be realised at European level of which 2 are cases of crisis responses. We assume that 9 cases would produce completely new statistical products and in 11 cases existing statistical products would be enhanced. Only for the latter cases, there would be burden reduction on the enterprise sector. New statistical products would be produced for all of the 5 crisis response cases. During the last 6 years, the European statistical system has explored and prepared a number of use cases for reuse of new data sources, such as reuse of mobile network metadata, financial (bank account and card) transactions, smart meter on energy and water consumption, real estate transactions, web platforms (accommodation, job advertisements, new forms of labour, mobility), mobility data related to different means of transport, smart personal data, smart farming data, environmental sensor data, which could be implemented depending on a cost benefit analysis and following the safeguards articulated in the policy measures of PO1.

The number of crisis response actions depends on the number of crises that will occur during the next decade. The current assumption is based on experiences from the current

crisis (Russian war of aggression against Ukraine and its effects on economy and society) and the COVID-19 crisis.

In addition, it is assumed that 4 cases of mandatory data sharing within the ESS would be implemented within 10 years after entry into force of the amended regulation. There are a certain number of existing statistical products, which could benefit from EU data sharing and for which there is a demand by users, such as balance of payments, foreign direct investments or migration.

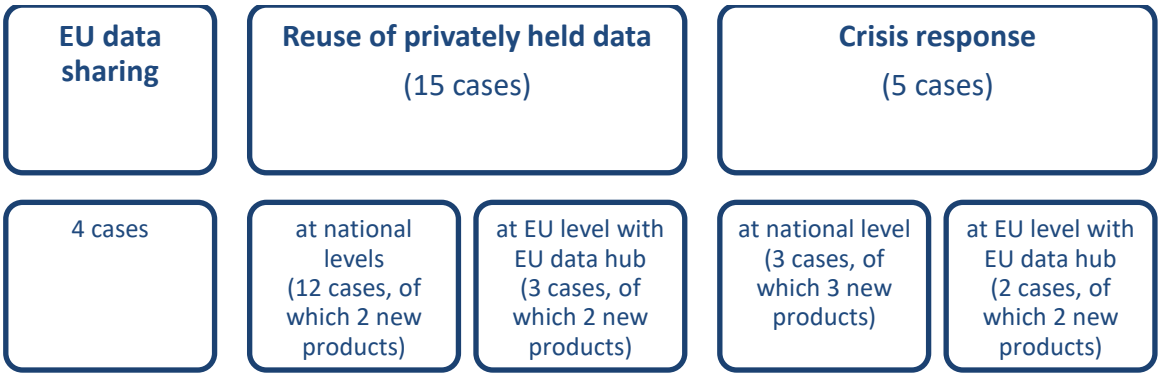

Figure 4: PO1: Application of statistical products / use cases

	A	B	C	D	E	F	G	H	I	J	K
1	Direct cost and benefit in Million EUR	B2G4S				Crisis				Total direct benefits and costs	
2		at national level		at EU level (Eurostat as hub)		at national level		at EU level (Eurostat as hub)		Total	
3		ESS	Businesses	ESS	Businesses	ESS	Businesses	ESS	Businesses	ESS	Businesses
4		12 Stat. products / domains		3 Stat. products / domains		3 Stat. products / domains		2 Stat. products / domains			
5	Preparation costs	30.6	30.6	0.6	0.6	1.6	1.6	0.0	0.0	32.9	32.9
6	Organisational	36.7	9.2	0.7	0.2	1.9	0.5	0.0	0.0	39.4	9.9
7	Meth. Development	165.2	18.4	3.2	0.4	8.7	1.0	0.2	0.0	177.4	19.7
8	Infrastructure	160.7	68.9	3.2	1.4	8.5	3.6	0.2	0.1	172.5	73.9
9	Operational	128.5	55.1	2.5	1.1	6.8	2.9	0.2	0.1	138.0	59.1
10	**Total costs**	**521.7**	**182.1**	**10.2**	**3.6**	**27.6**	**9.6**	**0.7**	**0.2**	**560.3**	**195.5**
11	Quality improvements	195.3		2.8		5.7		0.1		203.9	
12	Savings / burden red.	441.0	441.0	3.5	3.5					444.5	444.5
13	Additional outputs	259.2		9.2		32.4		0.8		301.6	
14	Compensation		15.3		0.3		0.8		0.0		16.4
15	**Total benefits**	**895.5**	**456.3**	**15.5**	**3.8**	**38.1**	**0.8**	**0.9**	**0.0**	**950.0**	**460.9**
16	**Net benefits / costs**	**373.8**	**274.2**	**5.3**	**0.2**	**10.4**	**-8.8**	**0.3**	**-0.2**	**389.7**	**265.4**

Table 11: PO1: Estimated differential cost and benefit for data reuse and crisis response[62]

[62] As compared with the baseline option (PO0)

The additional benefits for reuse of privately held data and the crisis response as compared to the baseline option outweigh the costs for both statistical offices and businesses as compared to the baseline option. The net benefits of EUR 389.7 Million (J16) for the statistical system and EUR 265.4 Million (K16) for the business sector are higher than for the baseline option due to more intensive use of the measures. The direct costs for the data holders for the cases producing new statistics exceed their benefits as there is no burden reduction effect for the business sector.

	A	B
1	ESS data sharing	
2	4 Stat. products / domains	Million EUR
3	Burden reduction	116
4	Savings through running a central system	23
5	Costs of running a central system	-0.7
6	**Total**	**138.3**

Table 12: PO1: Estimated differential benefits and costs of the measure on mandatory ESS data sharing[63]

I. Overview of Benefits (total for all provisions) – Preferred Option		
Description	***Amount***	***Comments***
Direct benefits		
Quality improvements of statistical outputs	204 million (Table 11, J11)	More granular statistics
Volume of statistics is increasing	302 million (Table 11, J13)	more statistical outputs
Increase in timeliness of statistics	-	not quantified, but estimated big effects in times of crises
More data available to research purposes	not quantified	
Central production of statistics leads to increased coherence	not quantified	
Indirect benefits		
Efficiency gains through better policy decision	not quantified	Society overall would benefit from direct European actions and data sharing due to better quality (granularity and timeliness) of statistics enabling better informed policy decisions.

[63] As compared with the baseline option (PO0)

Efficiency gains through improved data governance and stewardship	not quantified	More efficient data sharing and increased interoperability between data spaces leading to increased quality of statistics
Efficiency gains for businesses due to better informed economic decisions	not quantified	More statistical output and improvements in quality (time and granularity) can be used by businesses for taking informed decisions. All enterprises will benefit from this effect, especially SMEs as they will usually not be able to reuse data from new sources.
Administrative cost savings related to the 'one in, one out' approach		
Burden reduction on businesses due to mandatory data sharing	116 million (Table 12, B3)	Lower sample sizes result in reduction of burden on businesses and citizens
Savings for the ESS due to lower survey sizes	445 million (Table 11, J12)	
Burden reduction on businesses due to lower survey sizes	445 million (Table 11, K12)	Elimination of duplicate data collections across member States
Savings for the ESS due to running a central system induced by mandatory data sharing	23 million (Table 12, B4)	Data will be processed at central servers instead of national data processing. This type of cost savings is also included in B2G4S and urgent user demands. In these cases, savings are hypothetical as related systems are newly created. Cost efficiencies could be quantified in comparison to implementations in each Member State.

Table 13: Overview of benefits – preferred option[64]

II. Overview of costs – Preferred option					
		Businesses		**Statistical Offices**	
		One-off	**Recurrent**	**One-off**	**Recurrent**
B2G4S (national implementaitons)	Direct adjustment costs	30.6	18.4	30.6	165.2
	Direct administrative costs	-	133.1	-	325.9
B2G4S (European implementations)	Direct adjustment costs	0.6	0.4	0.6	3.2
	Direct administrative costs		2.6	-	6.4
Urgent Demand (national implementation)	Direct adjustment costs	1.6	1.0	1.6	8.7
	Direct administrative costs	-	7.0	-	17.3
Urgent Demand (European implementations)	Direct adjustment costs	0.0	0.0	0.0	0.2
	Direct administrative costs	-	0.2	-	0.4
Cost for mandatory data sharing	Direct administrative costs			2.4	1.4

[64] As compared with the baseline option

	Indirect costs	-	-	-	-
Costs related to the 'one in, one out' approach					
Total	Direct adjustment costs	32.9	162.7		
	Indirect adjustment costs				
	Administrative costs (for offsetting)				

Table 14: Overview of costs – preferred option[65]

The figures of Table 14 are calculated on the basis of the model that produces the figures in Table 11 by categorising them as adjustment or administrative costs. Costs related to preparations and methodological development are considered as adjustment while the other costs (organisational, operational and infrastructure) are categorized as administrative costs. The costs and benefits between statistical offices and the business sector are distributed according to the following Table 15.

		NSI / Eurostat	Data holder
Preparation costs	adjustment	0.5	0.5
Organisational	administrative	0.8	0.2
Meth. Development	adjustment	0.9	0.1
Infrastructure	administrative	0.7	0.3
Operational	administrative	0.7	0.3
Compensation	administrative	0	1
Quality improvements	Benefits	1	0
Savings / burden red.		0.5	0.5
Additional outputs		1	0

Table 15: PO1 – Distribution of costs and benefits between statistical offices and data holders

Policy option 2 (PO2)

For policy option 3, a total of 5 cases of reuse of privately held data (B2G4S) and 10 cases of urgent user data demands in times crises would be realised within the next 10 years. Out of the total of these 15 cases (5 B2G4S and 10 urgent user demands in times of crises), 6 would be realised at European level of which 4 are cases of urgent user demands in times of crises. We assume that 8 cases would produce completely new statistical products and in 7 cases existing statistical products would be enhanced. Only for the latter cases, there would be burden reduction on the enterprise sector. In contrast to PO1, the majority of use cases under PO2 relates to the policy measures that aim to address urgent user demands in times of crises. The increase is motivated by the latter mechanism. The number of 10 cases seems to be plausible based on recent experience with an increase in the demand of urgent user needs to provide statistics for high priority

[65] As compared with the baseline option

policy areas. This increase as compared to PO1 is motivated by measure 3.5, which allows initiating statistical actions in response to urgent user demands other than crises. In contrast to PO1, we would expect that data collections due to this mechanism would as well lead to improving existing statistical products having a temporary effect of burden reduction on businesses. We estimate the cases of reuse of privately held data considerably lower than for PO1 because of the missing consultation phase, which would lead to better quality of data reuse demands by statistical offices, and the missing dispute resolution mechanisms, which would lead to more court cases due to complaints by data holders. Moreover, businesses may be troubled by the lack of safeguards and incurrence of costs without compensation. There is evidence for such a likely development in the stakeholder interviews[66].

In addition, it is assumed that 6 cases of mandatory data sharing within the ESS would be implemented within 10 years after entry into force of the amended regulation. We assume a higher number of EU data sharing activities as this data sharing would also contribute to improved quality of European statistics based notably on coherence and comparability across countries.

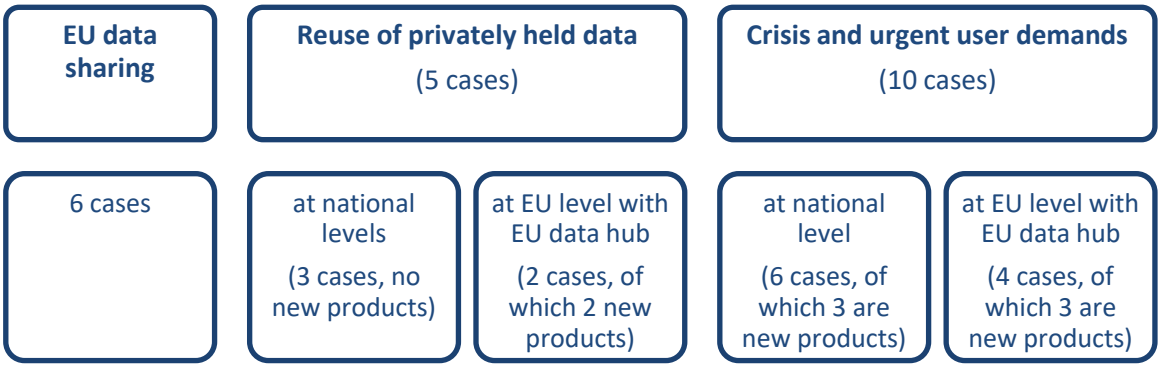

Figure 5: PO2 - Application of statistical products / use cases

	A	B	C	D	E	F	G	H	I	J	K
1	Direct cost and benefit in Million EUR	B2G4S				Crisis and urgent user demands				Total direct benefits and costs	
2		at national level		at EU level (Eurostat as hub)		at national level		at EU level (Eurostat as hub)		Total	
3		ESS	Businesses	ESS	Businesses	ESS	Businesses	ESS	Businesses	ESS	Businesses
4		3 Stat. products / domains		2 Stat. products / domains		6 Stat. products / domains		4 Stat. products / domains			
5	Preparation costs	6.3	6.3	0.4	0.4	3.2	3.2	0.0	0.0	10.0	10.0
6	Organisational	6.3	3.1	0.4	0.2	3.4	1.5	0.0	0.0	10.2	4.8
7	Meth. Development	34.0	3.8	2.2	0.2	17.5	1.9	0.1	0.0	53.8	6.0
8	Infrastructure	27.0	20.3	1.8	1.2	14.6	9.7	0.1	0.1	43.5	31.3
9	Operational	24.0	17.8	1.6	1.0	12.2	8.1	0.1	0.1	37.8	27.0

[66] This development can be observed in the case of the new Norwegian statistics regulation, which misses this mechanism.

14	None-compensation						0.8				0.8
10	**Total costs**	**97.7**	**51.3**	**6.3**	**3.1**	**50.9**	**25.3**	**0.3**	**0.2**	**155.2**	**79.8**
11	Quality improvements	44.1		1.4		17.0		0.2		62.7	
12	Savings / burden red.	138.6	81.9			34.0	22.7	0.8	0.6	173.5	105.1
13	Additional outputs	37.8		8.0		42.1		-0.1		87.8	
15	**Total benefits**	**220.5**	**81.9**	**9.4**		**93.2**	**22.7**	**1.0**	**0.6**	**324.0**	**105.1**
16	**Net benefits / costs**	**122.8**	**30.6**	**3.1**	**-3.1**	**42.3**	**-2.6**	**0.7**	**0.3**	**168.8**	**25.3**

Table 16: PO2 – estimated differential costs and benefits for data reuse and crises response[67]

The additional total net benefits of PO2 for the cases of reuse of privately held data and responses to user demands in times of crises as compared to the baseline option are much lower (168.8 Million (J16) for statistical offices and 25.3 Million (K16) for businesses) than those of PO1. This is mainly due to the lower number of cases. In addition, the relatively higher number of cases in which new statistical output is produced does result in lower burden reduction of the business sector. Finally, the data holders are in this scenario not compensated for any cost related to making data available for reuse.

		NSI / ESTAT	Data holder
Preparation costs	adjustment	0.5	0.5
Organisational	administrative	0.7	0.3
Meth. Development	adjustment	0.9	0.1
Infrastructure	administrative	0.6	0.4
Operational	administrative	0.6	0.4
Additional outputs		1	0
Quality improvements	Benefits	0.6	0.4
Savings / burden red.		1	0

Table 17: PO2 – Distribution of costs and benefits between statistical offices and data holders

In PO2, the distribution of cost and benefits puts a higher share of cost on data holders due to additional obligations on the data holders, such as the appointment of data stewards, additional requirements regarding metadata and quality documentation, and the lack of consultations. At the same time, statistical offices would have higher benefits from improved data documentation. Finally, all costs of data access have to be carried by the data holders. Compensation mechanisms are not foreseen.

	A	B

[67] As compared with the baseline option (PO0)

1	ESS data sharing	
2	6 Stat. products / domains	Million EUR
3	Burden reduction	151
4	Savings through running a central system	27
5	Costs of running a central system	-1.3
6	**Total**	**176.7**

Table 18: PO2 - Estimated differential benefits of the measure on mandatory ESS data sharing[68]

Due to the mandatory of participation in ESS data sharing activities, we assume 50% more cases of data sharing. However, we assume a lesser increase in savings due to burden reduction, which are based on the assumption that mirror data collections can be avoided or at least reduced. We assume that the ESS would concentrate on those cases with highest burden reduction. Hence, the more data sharing cases are implemented the lower are the efficiency gains per additional case.

PO2 contains a measure providing access to the data shared by data holders for research purposes. This is an additional burden on statistical offices, which would anyhow process micro data, which is used to produce aggregate statistics. The following Table 19 shows the additional cost for statistical offices due to this measure. It is assumed that the number of access requests is 50 per year per statistical office for these kinds of data. The assumptions are motivated through experience of Eurostat executing this service for traditional statistics. The estimated access requests of this category are 12.5% of the actual total of annual requests[69].

Access to microdata for research (1000 EUR)	**NSIs**	**Eurostat**	**Total**
Staff cost	1283	48	**1330**
Infrastructure cost	675	25	**700**
Other cost	192	7	**200**
Total	**2150**	**80**	**2230**

Table 19: PO2 - Estimated costs of additionally providing access to shared raw data for research purposes

[68] As compared with the baseline option (PO0)

[69] The assumptions are motivated by the figures in the impact assessment for the proposal of the Data Act and by experiences of Eurostat related to supplying a service of data access to confidential data for research.
European Commission, impact assessment Report accompanying the Proposal for a Regulation on harmonised rules on fair access to and use of data (Data Act), SWD(2022) 34 final, page 106

3. RELEVANT SUSTAINABLE DEVELOPMENT GOALS

The preferred PO1 covers various measures to improve the EU level data evidence potentially relevant for several, or even all, sustainable development goals (SDGs). However, impacts in terms of actual progress towards any of the SDGs stemming from improved data evidence are naturally indirect and hard to assess given that the initiative at hand is a framework regulation intended to improve the data availability, responsiveness and data sharing generally, which will only manifest itself in sectoral legislation. For instance, if used for improved SDG policy-making, actual progress would largely depend on dispersed contextual factors and on the sensitivity of policy impacts on the availability and quality of data evidence. Based on the use cases in Annex 6, Table 20 below identifies SDGs that could potentially benefit from this initiative.

Table 20: Overview of relevant Sustainable Development Goals – Preferred Option

III. Overview of relevant Sustainable Development Goals – Preferred Option(s)		
Relevant SDG	**Expected progress towards the Goal**	**Comments**
1 – No poverty		• Improving data from farms and the agricultural sector
8 – Decent work and economic growth		• Improving data the labour force, data on income and living conditions and data from national accounts.
9 – Industry, innovation and infrastructure		• Improving statistics regarding goods, services and foreign direct investments, as well as regarding business statistics and digital trade and cross-border data flows
11 – Sustainable cities and communities		• integrating geospatial data and statistical quantitative data, mostly at a regional and subregional level. • improving regional or urban statistics. • improving population grids of importance for policy areas such as transport, mobility or the environment.
17 – Partnership for the goals		• increasing availability of sound data and measures to promote the accountability, which can serve as a model

ANNEX 4: ANALYTICAL METHODS

Regulation (EC) No 223/2009 is a framework regulation, the revision of which will have its main effect through subsequent other legislation, at both the EU and national level. The analytical methods of assessing the impact of the revision take this into account. Not only will the revision have elements that need to be worked out subsequently (see chapter 8), but it will also not identify the data sources that will actually be reused for statistical purposes. That will be done through the Annual Work Programme for European statistics (AWP). Every time a statistic is added to the AWP that makes use of privately held data, either as its main source or as a supplementary source, this will require a convincing justification, of which an assessment of the costs and benefits to society, with a breakdown to groups of stakeholders, will be part. When executing the AWP, specific data holders will be identified and requested to enable the reuse of the data held by them. Such requests must also be explicitly justified.

For assessing the impact of the revision under these conditions, a use case approach has been taken. Based on earlier research (see Annex 1.4), a number of use cases of potential use of new data sources for official statistics were derived. This was already done in the context of the impact assessment for the Data Act. For some use cases, it was possible to make an assessment of the associated costs and benefits. Depending on assumptions about the eventual extent of using new data sources in official statistics and their similarity to use cases already assessed, total costs and benefits for regulating the use of new data sources for official statistics could be estimated, albeit assumption-based and with a large margin of error (see Annex 3). The use cases that fed into the impact assessment for the Data Act were also input to the study carried out by ICF, which used them, together with additional use cases, for their in-depth interviews with stakeholders (see Annex 2), to get a more concrete view on the obstacles, the potential costs to data holders and producers of statistics, and the potential benefits for, among other things, policymaking. The remainder of this annex describes the analytical methods of the ICF study and the analytical approach taken to the assessment of costs and benefits of Annex 3, i.e., of what has been used from the impact assessment for the Data Act.

Analytical methods of the ICF study

The analytical methods used by ICF in their study are described in [ICF, Annex 5]. All the data collection tools in that study were designed to collect both quantitative and qualitative evidence. In particular, questions for relevant cost items were included in the online survey to aim for a full standard cost model (SCM). An SCM is achieved by estimating the impact on costs on the relevant population of relevant stakeholders. This impact corresponds to the formula PQT: Price (in EUR of the additional cost/burden) * Quantity (population of products, people, etc.) * Time (one-off or recurring). This modelling approach requires extensive granular data on the specific segment of the population and/or business sectors, monetary values of costs or additional FTEs. The ICF methodology was designed to follow the traditional steps of an SCM:

1. identifying the additional burdens/costs;
2. identifying target group and/or sector(s);
3. identifying the frequency of costs and parameters;
4. design survey to gather all the aspects in steps 1 to 3;
5. collect data and prepare template for the analysis and reporting;

6. reporting.

However, the data gathered in the online survey were not sufficient to feed the model to estimate the costs. Although the questionnaire included questions on quantitative impacts to allow for quantification, respondents were mostly unable to provide quantitative figures. It is known that some of the challenges related to the cost and benefit analysis usually include:

- Small sample size for robust estimations;
- Stakeholders unable to provide coherent quantifications of cost items;
- Selection bias of responses;
- Difficulty in establishing causal links between costs and benefits because of the broad nature of policy options;
- Consistency across results with testable assumptions. For example, if the sample is skewed towards a specific category of stakeholders, should the first item in the list render a re-sampling impossible, then generalising assumptions would be unfeasible.

While the ICF study was designed to avoid these issues, they still emerged during the data collection. Interviews for use cases provided some quantitative estimates of costs and benefits, but hardly generalizable as they were dependent on use cases' characteristics, namely geography, stakeholders involved, timing, digital infrastructures and business processes. The targeted survey was designed to gather both quantitative and qualitative estimates, but stakeholders again could not directly provide quantitative figures on costs and benefits. While respondents could not quantify, they provided a systematic assessment (using a 1-10 Likert scale) of impacts, which was used for the qualitative appraisal. Moreover, the small sample size (N=27) would have made any quantitative estimate not entirely robust. Therefore, costs and benefits for this study had to rely on qualitative evidence, including and especially the anecdotal figures coming from the use cases.

An important example of qualitative cost and benefit information is the use case on financial transactions data. In this use case, the Norwegian NSI has an ongoing data sharing activity with a company called Vipps. The aim is to collect debit card transactions data with the objective of gathering information about the retail sector in Norway. The cost-benefit analysis conducted by the Norwegian NSI[70] found that despite the total costs for the private data holder in terms of development and data delivery (1 or 2 months of full-time work), the data sharing would reduce the overall burden on the Norwegian business sector by up to 3.400 hours of work, equalling almost 23 months of cumulative work. This example clearly shows that, in this specific use case, the benefits of mandatory data sharing (for all stakeholders, including the business sector, when considered in its entirety) clearly outweigh the costs.

Analytical methods of the costs and benefits analysis used as input to the impact assessment for the Data Act

[70] Norway is one of the few countries that has a statistics act that mandates data collection from private data holders. According to the Norwegian Statistics Act, the Norwegian NSI needs to conduct a cost-benefit analysis before compelling private data holders to deliver data.

The study on "Methodological support to impact assessment of using privately held data by official statistics" mentioned earlier[71] offered an indicative example of data sharing with MNO data in the context of Covid-19 and scanner data[72]. It argued that reusing privately held data by public producers of official statistics yields benefits for NSIs, for society in general and for the business community.

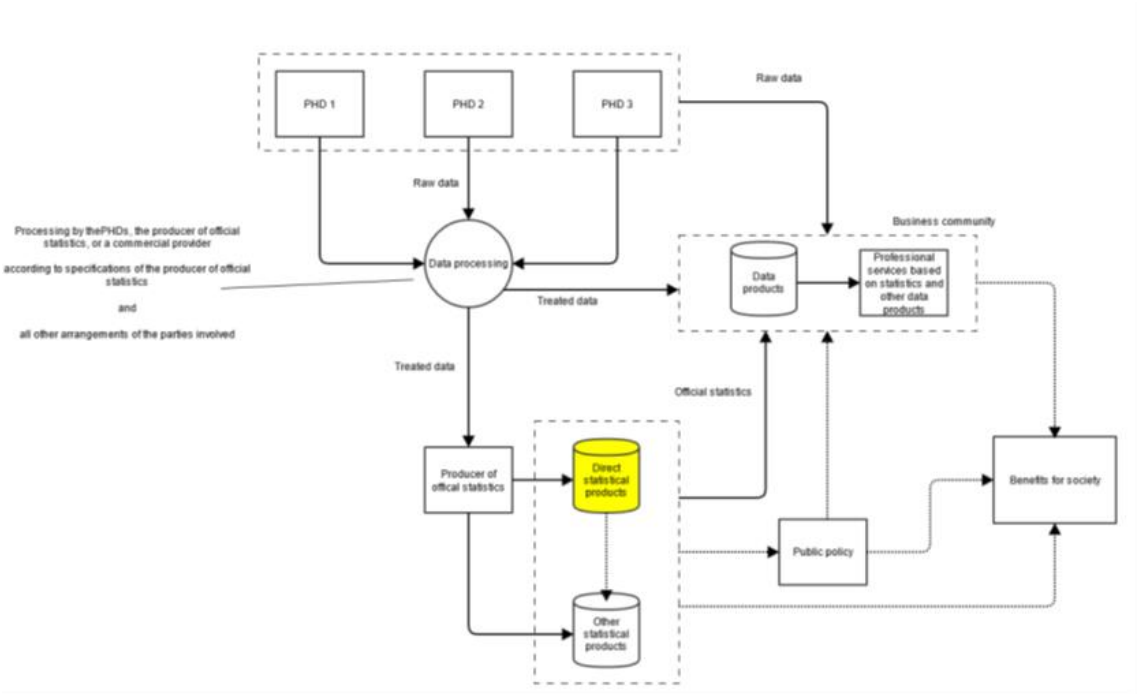

Figure 6: Visual representation of B2G data sharing chains for official statistical production, and of the benefits and the B2B data sharing chains they generate

The study stated that smarter lockdown measures would result from reusing mobility data coming from MNOs: "If the use of MNO data could make things efficient and save one week of work it would reduce the damage by roughly 0.2% of GDP, which is substantial." Moreover, if an NSI could institute a long-term established process with recurring outputs, costs can be as low as EUR 1 million, "which would be a tiny and insignificant fraction compared to the benefits – particularly under COVID conditions" (p23). For scanner data, the study estimated that the balance of benefits and costs would even be higher.

The study also proposed two alternative methods for quantifying B2G4S: bottom-up and top down. Using the bottom-up approach and the above estimates on scanner data[73], the study suggested that this type of data sharing for 12 data sources in each of the 27 Member States could add, in total, approximately 6.7% to the data economy, which is estimated at 3.5% of GDP. Employing a bottom-down approach, the study calculated that if the amount of official statistics is increased by 20% because of B2G data sharing, this

[71] Sciadas G and Stavropoulos P., Methodological support to impact assessment of using privately held data by official statistics, Literature review and model, December 2021.
[72] It also pointed out that "conclusive modelling of costs and benefits will undoubtedly have to wait for sufficient practical experiences to accumulate" (p3).
[73] As well as the open data literature, which indicates that induced benefits can be higher than direct benefits by a factor between 20-50 times.

will bring an added value to the economy of EU27 + UK of between EUR 4.3 billion (20% of 21.5) and EUR 12.3 billion (20% of 61.3) until 2030.

To put the quantification in this impact assessment in perspective, a few other, broader studies can me mentioned. The "Study to support an impact assessment on enhancing the use of data in Europe" concluded that the economic impact of the policy measures to foster data reuse as compared to the baseline scenario could imply an increase in GDP with 273 billion EUR (representing an additional 1.98% of GDP)[74]. Another study analysed the size of the data market assuming data products created from processing and reusing data are exchanged. The impact of the data economy is considered the sum of direct, indirect and induced impacts. The study estimated an overall impact of the data economy on the EU27 GDP of 2.1% (EUR 243 billion) in 2016 and 2.2% (EUR 267 billion) in 2017. Further projections stated that, by 2025 and under certain assumptions, it could grow as high as 5.4% of GDP[75].

ANNEX 5: LEGAL CONTEXT

The revision of Regulation (EC) No 223/2009 takes place in the following legal context.

The **European Data Strategy** put forward by the Commission in February 2020[76] aims to speed up the development of the European economy and to harness the value of data for the benefit of the European society. It provides the overall policy context under which the Regulation (EC) No 223/2009 has to be analysed and the fitness of the European statistics for the digital age has to be assessed.

The **proposal for a Data Act**[77] adopted by the Commission on 23 February 2022 is one of the main pillars of the European Data Strategy. Its main objective is to set the right conditions for better control and conditions for data sharing for citizens and businesses. The Data Act covers both business-to-business (B2B) and business-to-government (B2G) data sharing and sets up rules regarding the use of data generated by Internet of Things (IoT) devices. The Data Act proposal introduces the possibility for public sector bodies to access and use data held by the private sector that is necessary for specific public interest purposes but under strict conditions and only in cases of emergencies and other exceptional needs.

The Data Act proposal follows essentially an internal market logic (as illustrated by its legal basis i.e., Article 114 TFEU, whose objective is the establishment and functioning of the internal market) and clearly states that it is without prejudice to rules addressing needs specific to individual sectors or areas of public interest and that it does not affect obligations laid down in Union or national law for the purposes of reporting, complying

[74] Deloitte, The Lisbon Council, JIIP, GovLab, Timelex, Odi (2022) "Study to support an impact assessment on enhancing the use of data in Europe".
[75] The European data market study update
[76] See https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_en
[77] Proposal for a Regulation of the European Parliament and of the Council on harmonised rules on fair access to and use of data (Data Act).

with information requests or demonstrating or verifying compliance with legal obligations.

The **Data Governance Act (DGA)**[78] is another important pillar of the European Data Strategy and will be applicable from September 2023. It is a cross-sectoral instrument that aims to promote the availability of data and build a trustworthy environment for data sharing to facilitate the use of data for the creation of innovative new services and products as well as research. It also aims to encourage the development of technical solutions, such as anonymization, pseudonymization or accessing data in secure processing environments (e.g., data rooms) supervised by the public sector. In that context, a number of 'common European data spaces' will be established, such as for health, the Green Deal, and energy. They will facilitate data pooling and sharing by creating infrastructure and governance frameworks to promote data-driven innovation. The DGA also establishes a European Data Innovation Board, which will address issues such as data interoperability between data-sharing organisations. Creating the necessary tools and processes for data sharing could be expected to benefit data sharing regardless of its direction, whether it is G2B or B2G.

The **Interoperable Europe Act** proposal[79] recently adopted by the Commission, is an act which aims to strengthen cross-border interoperability and cooperation in the public sector across the EU. The goal of the Act is to move beyond the current voluntary cooperation on interoperability in the EU by establishing a structured EU cooperation where public administrations, supported by public and private actors, come together in the framework of projects. The Act also establishes mandatory interoperability assessments to evaluate the impact of changes in information technology systems on cross-border interoperability in the EU, foresees a revision of the Europe Interoperability Framework and will promote share and reuse of solutions, often open source, through an Interoperable Europe Portal.

The **INSPIRE Directive**[80], establishing an infrastructure for spatial information in Europe to support Community environmental policies, and policies or activities which may have an impact on the environment, entered into force in May 2007 with a view to ensuring that the spatial data infrastructures of the Member States are compatible and usable in a Community and transboundary context. The Directive requires that common Implementing Rules (IR) are adopted in a number of specific areas (Metadata, Data Specifications, Network Services, Data and Service Sharing and Monitoring and Reporting).

The ongoing discussions on the **proposal for an ePrivacy Regulation** are also relevant for this initiative. One of the main objectives of the proposal is to guarantee privacy rules for new players providing electronic communications but also for communications content and metadata. Additionally, the Regulation is expected to foster innovation as

---

[78] Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724.

[79] Commission proposal for a Regulation of the European Parliament and of the Council laying down measures for a high level of public sector interoperability across the Union (Interoperable Europe Act), COM(2022) 720 final, adopted on 18 November 2022.

[80] Cf. Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE), OJ L *108, 25.4.2007, p. 1*.

once consent is given for communications data to be processed, traditional telecoms operators will have more opportunities to provide additional services and to develop their businesses. This could support the use of data analytics and data sharing for the purpose of the production of official statistics as well.

With the **General Data Protection Regulatio**n (GDPR)[81] and the **ePrivacy Directive**[82], the EU has put in place a solid and trusted legal framework for the protection of personal data. More specifically, Article 5(1), point (b), of the GDPR provides that the further processing of personal data for scientific or historical research purposes or statistical purposes should, in accordance with Article 89(1) of Regulation (EU) 2016/679, not be considered to be incompatible with the initial purposes.

The revision of Regulation (EC) No 223/2009 should ensure that any B2G data-sharing partnership for the sake of compiling official statistics that will include personal data shall be carried out in full compliance with GDPR. Namely that individuals will continue to be empowered to take control of their data by exercising their rights under the GDPR, namely giving and revoking consent, requesting data access and erasure of data and requesting data portability. In 2018, the European Commission published a set of principles on B2G data sharing in its communication 'Towards a common European data space'[83] and accompanying staff working document 'Guidance on sharing private-sector data in the European Data Economy'[84]. These documents defined six principles that could support the supply of private-sector data to public-sector bodies under preferential conditions for reuse, such as, proportionality, purpose limitation, transparency and accountability. These principles are aligned with the ones provided by the GDPR.

Finally, the **Single Market Emergency Initiative (SMEI)**[85] aims to put in place a flexible and transparent mechanism to respond quickly to emergencies and crises that threaten the functioning of the single market. It aims to ensure the coordination, solidarity and coherence of the EU crisis response and protect the single market's functioning. The instrument would complement other policy tools to anticipate and prevent disruptions, where possible, and would also prepare for and respond to unavoidable crises, which have important cross-border effects and threaten the functioning of the Single Market. The proposal currently foresees the possibility for the Commission to invite or request under certain conditions, targeted information from the economic operators in crisis-relevant supply chains, for the management of the SMEI or for compiling relevant official statistics.

---

[81] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).

[82] Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector.

[83] See https://digital-strategy.ec.europa.eu/en/news/communication-towards-common-european-data-space

[84] See https://digital-strategy.ec.europa.eu/en/policies/private-sector-data-sharing

[85] See Proposal for a Regulation of the European Parliament and of the Council establishing a Single Market emergency instrument and repealing Council Regulation No (EC) 2679/98.

ANNEX 6: ASSESSED USE CASES

Introduction

This annex describes the results of extensive in-depth interviews by ICF with various stakeholders: mainly private data holders, institutional users of statistics, and NSIs. With each of the data sources limited explorative experiments had been carried out in partnerships between data holders and NSIs. The assessment includes the following five digital data sources:

1. data from Mobile Network Operators (MNOs)
2. financial transactions data
3. data from smart meters
4. data from smart devices (mainly from the IoT), and
5. data from collaborative economy platforms

Before summarising the results, it is important to understand the relationship between the role of various categories of stakeholders and costs and benefits. Figure 5 provides a model of stakeholders [ICF, section 2.5]:

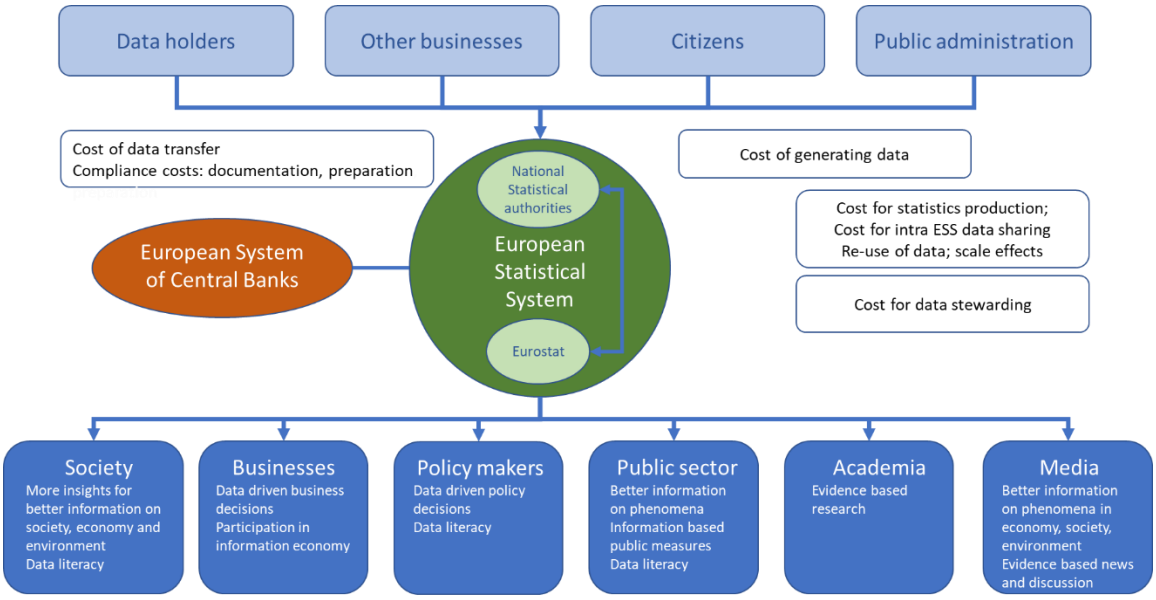

Figure 7: Model of stakeholders related to the ESS [source: ICF]

Benefits and costs are incurred by different types of stakeholders and relate to actions taken by those stakeholders (not necessarily the same ones). For European statistics, the process starts with the private holders of data (Figure 5, top left), who incur costs for complying with the requests from the statistical office; this interaction may result in benefits as well. The national statistical authority produces statistics, which incurs costs. The users of European statistics a heterogeneous group, of which every user can enjoy direct benefits and incur costs by making use of these statistics. These uses, in particular to the extent that they influence or even shape policies, will influence all stakeholders in

society that are subject of those policies, leading to (indirect) benefits and costs pertaining to those stakeholders.

Summary of the results of the five use cases

Types of data and scope

The use cases provided valuable examples on the type of new digital data sources that can be re-used to generate official statistics:

- MNOs: network signalling data were used to understand population distribution across the national territory in the context of Covid-19; MNO's data were also used to produce mobility statistics on commuting and tourism;
- Financial transaction data: debit card transactions were used to produce statistical indicators on the retail sector in one country;
- Smart meters: energy data coming from smart meters informed the total amount of electricity being consumed and, when combined with other sources, provided insights into new residential construction, the energy efficiency of buildings; they were also used to better understand where more work is needed to install meters and to move residential dwellings away from solid fuel use.
- Smart devices (IoT): traffic, air quality, and noise data gathered from IoT devices were used to understand levels of pollution within and across regions;
- Collaborative economy platforms (tourism): accommodation bookings aggregated at the city level were used to better understand trends in tourism.

Operational implementation

There were several themes emerging from the five use cases on the operational implementation of these data sharing activities:

- The heterogeneous nature of these B2G4S activities makes it difficult to highlight commonalities in their operational implementation. An assessment of each use case separately was provided by ICF [ICF, chapter 6].
- Different legal frameworks were used to implement B2G4S activities, including specific public contracts, mandatory data delivery based on a national statistics law, a cooperation agreement and a non-disclosure agreement. This extreme variety of legal frameworks reiterates the above point about the current heterogeneity of data sharing activities, reinforcing the argument on the need of further harmonization in the EU.
- The level of effort required to implement the data sharing activities was dependent on the novelty of the collaboration: in one use case (financial transaction data), the operational implementation required less efforts because stakeholders had already previously collaborated and had already the technical infrastructure in place; in another use case (MNOs), the data sharing required the private data holders to develop a series of algorithms with a guidance from the NSI. One smart meter use case showed that it could take up to five years for data sharing collaborations to be established and ensure they meet all the legal obligations and safeguards (especially in relation to data protection).
- In one use case (financial transactions data), the NSI needed to produce a cost-benefit analysis before persuading the company to deliver the data. This analysis had to show evidence that benefits of data sharing outweighed its costs.

- Stakeholders always put strong emphasis on secure data sharing and data anonymity (to ensure users' data were not identifiable), suggesting these are key factors in the implementation of any data sharing activity. For NSIs this is already standard practice.
- In at least two use cases, the data sharing agreement was facilitated by Eurostat, highlighting the potential role of the institution in coordinating any future data sharing activity with the objective of understanding cross-border EU phenomena.

Incentives and benefits

Use cases underlined that all main stakeholders (statistical users and society, NSIs and private data holders) can benefit from B2G4S data sharing, although in different ways.

Overall, stakeholders found it difficult to quantify benefits, especially when considering more indirect ones such as improved decision-making. However, one useful indicator could be the number of direct data collections (administrative forms and/or surveys) directed at business and households that can be replaced by data sharing collaboration.

Statistical users

When new data sharing activities were able to fill current knowledge gaps, statistical users improved their decision-making, ultimately designing better policies for the benefit of society and the public good.

For example, data sharing with online tourism platforms can help understand how travelling is resuming post Covid-19, gaining a more solid knowledge of the impact of tourism on local, regional, national and European levels.

In another example on mobile network data, mobility indicators generated with MNO data were extremely important to better understand interactions between territories and the daily behaviours of individuals, which helped informing public policies on spatial planning and land use.

NSIs

When data sharing activities were built on a solid foundation, they brought clear benefits in terms of quantity, quality, granularity, frequency and timeliness of data. For example, one NSI assured that these data allow to go far beyond what would be possible to achieve through traditional data gathering methods such as surveys, primarily due to their connected and digital aspect. In addition to faster access, the data is in some cases also more accurate, moving away from estimations and customer submitted entries. This can increase the overall reputation of NSIs, if this means an increased capacity to fill current information gaps.

Furthermore, (automated) data sharing can also reduce the burden on NSIs, businesses and society if these can replace filling lengthy paper or online surveys. For example, one NSI was able to cut up to 20,000 individual questionnaires per year, each of them taking approximately 10 minutes to deliver. This would reduce response burden for the business sector by up to 3.400 hours.

Private data holders

Interactions with competent statistical officers can improve companies' internal methodologies and overall data quality, which might lead to the possibility of using these data for commercial purposes. In one use case (financial transactions data), the data sharing was also considered to benefit the company's commercial interest as being able to tell their (paying) clients that they are working with Eurostat (or other public sector authorities) brings some "gravitas" to their dataset.

In one use case, MNOs were compensated, but this case was unique among all other case studies. In general, there is disagreement among stakeholders regarding what the right approach should be regarding financial incentives.

Other important benefits may include compliance with national laws (when data sharing is already mandatory) and participation to research projects.

Companies' social corporate responsibility managers could also rightly claim their companies are contributing to improving societies by offering data which help producing better policies, which ultimately strongly enhance their reputation.

Costs

As for the costs, quantification and generalisation is hard, mainly because there are multiple factors to consider when designing and implementing a data sharing collaboration. For example, data sharing activities between entities that have already collaborated (hence with already established procedures, infrastructures, and dedicated staff) will have much lower costs compared to a data sharing collaboration starting from scratch.

While stakeholders found it difficult to quantify costs, one could use as an indicator the number of work hours each employee logged in setting up and implementing the data sharing activity. Some use cases presented anecdotal evidence going in this direction.

Use cases generally featured examples of data collaborations where all parties covered their own costs. However, in one case private data holders were compensated for their data processing and analysis services: a three-year contract was stipulated for a value of EUR 300.000, with subsequent annual contracts worth EUR 24.000. Many NSIs are sceptical about a model in which they would have to pay to reuse privately held data. The ESS is in favour of introducing "adequate mechanisms and incentives to deal with the initial investments required and marginal costs that might be incurred by the data holders in processing, especially aggregating or running algorithms on the primary data to make them ready for use in official statistics"[86].

NSIs generally agreed that costs on their side would be low or manageable, even though this might vary according to the level of ambition of the data sharing activity. For example, one NSI (financial data transactions) needed a team of 5/6 people for 2/3 months to set up the data sharing activity, but this was possible because the collaboration (with the same company) had already taken place in the past. Another NSI using smart meter data reiterated that personnel costs were minimal in comparison to collecting data

[86] ESS position paper on the future Data Act proposal.

in a traditional way. On the other hand, still another NSI noted that a lack of existing competences and skills within the organisation meant they had to hire a data scientist. Finally, NSIs need to have the appropriate resources to handle the data coming from private data holders as there is significant pre-processing and cleaning before data are useful for statistical purposes.

Private data holders could incur investment, development and data delivery costs, which were anecdotally and occasionally "quantified" in terms of full-time employees involved in the data sharing. One company estimated the total costs to be equal to 1 or 2 months of full-time work, but this is clearly influenced by the fact that the collaboration (with the same NSI) had already taken place in the past. Interviewed private holders usually thought costs were low as well. One private data holder in the economy platform use case stated that costs are manageable if there is one data portal, one API, and one format. Nonetheless, they could become unmanageable if the data holder had to share data with each Member State or share it at very short intervals. Generally, use cases showed that there are companies which are "readier" (have the technological know-how, the infrastructure, the personnel etc.) to engage in data sharing. For these companies, it would be easier to start and sustain a data sharing compared to other companies which are less tech savvy.

Challenges

Use cases highlighted several challenges that need to be overcome to implement successful data sharing collaborations:

- The absence of a clear legal framework, coupled with the complexity of national legislative systems, were considered obstacles for effective data sharing. One NSI wanting to access smart thermostat data was unable to overcome the legal issues in access. Whilst the private company was keen to collaborate, the absence of a clause within the customer contract detailing the sharing of data for statistical purposes, made data access impossible.
- Data privacy and security were the most pressing issues across all use cases, with private data holders often referring to privacy obligations towards their customers as a reason to start or end data collaborations. Companies wanted to avoid giving the impression that they establish data sharing activities with public authorities for the purpose of tracking individuals. The reputational risk they face if data were leaked is often seen as very high. Depending on the nature of the data sharing activity, other potential legal issues highlighted were related to data ownership and copyright.
- Commercial interest was another frequently cited issue, with companies fearing to lose their competitive edge when data is shared with a public authority, if 1) their technical innovations are revealed/leaked in the process, 2) they already sell data commercially. Interestingly, one company in the financial transactions data use case suggested that their commercial interest is not hurt as the data they sell to their clients serve very different purposes.
- The compatibility of concepts, measures, and methods could become a big obstacle if not harmonized among stakeholders. If data collection definitions and methodologies are not transparent, official statistics cannot be produced.
- Data quality could become an issue when the shared data are not known to be representative of the whole population; moreover, these types of data often lack

socio-demographic context (as opposed to surveys, which can collect those contextual data) which limit their potential applications.

- One-off or limited-in-time data sharing activities were seen as problematic by NSIs, which often require time to ensure that the production of official statistics comply with the appropriate standards and regulations. On the other hand, private data holders in the mobile network data use case have expressed concerns against the prospect of other public authorities asking them for data, which over time would create additional burdens on their internal work.

ANNEX 7: DATA GAPS IN TERMS OF CONTENT, FREQUENCY, TIMILINESS AND GRANULARITY OF EUROPEAN STATISTICS

This annex provides a summary of the gaps identified in European statistics. This is based on the extensive in-depth interviews by ICF with various institutional users of statistics. Certain direct effects of the gaps are also summarised per area of policymaking as well as instances of potential uses of private data sources that could help closing these gaps and their perceived benefits.

Identified gaps and their effects

Certain gaps could be identified in the statistical needs of all users consulted and these generally follow the next trends:

- Statistics on labour mobility or employment on digital platforms, which are necessary for decision-making in the area of employment are insufficient. The need for this type of data on a more regular basis was highlighted. Here, for some of the existing statistical series (e.g., SILC), the biggest issue reported was also timeliness of the data. A scarcity of available breakdowns was also reported as an issue especially in the case of policy areas related to migration, labour mobility and the digital economy.

Perceived effects: Recent crises have highlighted a need for more frequent monitoring as the social situation is quickly developing. Policy decisions need to be taken based on information, which is accurate. For example, social statistics inform the European Semester reports and country specific recommendations, and their accuracy depends on the evidence available.

- Statistics on public transport is another major data gap. A daytime population grid would be very useful to understand where people spend most of their time during the day. That will inform flows, exposures to risks to pollution, etc. and will also give a lot of information on transport needs, transport flows. Good data on residential energy consumption is also lacking.

Perceived effects: Funding is generally targeted towards less developed regions or Member States, and unfortunately, less developed areas are typically the ones where data is often entirely missing. For policy makers, it would be much easier to target support to areas where it is needed if they could measure the current offer. Moreover, more funding is needed to provide alternatives, to fund household surveys, general population surveys

or enterprise surveys, or to buy commercial data. This kind of surveys can amount to costs ranging from 3 to 5 million euros a year. Due to lack of data on residential energy consumption, investments in energy efficiency are done in a relatively blind manner. The cohesion policy directs a budget of approx. €350 billion, which is one third of the EU budget as a whole, so the effects of inappropriate policy decisions would be extremely costly. Conversely, even if it is just a marginal improvement of 1%, when compared to such budgets, the benefits would be a big consideration.

- Another major issue is access to disaggregated data, especially at micro level as for market monitoring and crisis management, this is extremely important. Timeliness of the data is also often reported as an issue.

Effects perceived: When evidence and underpinning data is missing, the most direct effect is that policymakers cannot fully show or demonstrate why certain policy options are the best to pursue. Similarly, when the data is too aggregated it's difficult to know what the policy will generate at a micro level. Delays of one to two years that sometimes occur have a significant effect on policymaking which relies heavily on markets monitoring and crisis management. This has also been identified in a recent audit by the European Court of Auditors[87], where they have been looking at the use of big data in agriculture. Their report concluded that the European Commission has not capitalised on the potential of big data for analysing and subsequently designing the EU's Common Agricultural Policy. As a result, the Court considers the Commission as not having enough evidence to comprehensively assess the CAP's needs and impact. (The CAP accounts for more than a third of the EU's budget – €408 billion between 2014 and 2020.)

The extra time needed to find alternative data sources is also an important downside of existing gaps in statistics. Another effect is the need to make significant investments in procuring commercial data. However, procurement is often unstable because of terminating or changing contracts, impinging on business continuity and the analytical work.

- Certain data on tourism are often missing such as same-day tourism, data collections on the destination level, local conditions, the satisfaction of the residents, and other types of aspects at the local level.

Effects perceived: Crises such as the Russian aggression in Ukraine increased the need to understand how this impacted tourism and data was not available as quickly as needed. The alternative was data provided by transport companies which helped to get an understanding given that there was no official data on both the concrete impacts on the travelling and tourism. Nevertheless, unofficial sources limit what policy makers can officially do to inform the public, their decision-making processes and the knowledge basis.

- The frequency at which labour market statistics or data on employment are provided sometimes impacts policymaking, especially in a crisis environment.

[87] See https://www.eca.europa.eu/en/Pages/NewsItem.aspx?nid=16713

Effects perceived: the analysis done by financial and economic institutions is less precise. This can lead to worse monetary policy decisions or decisions in the area of financial stability.

- More data on market concentration and on market power is needed, more granular information on price levels for products and services, and more statistical data on market dynamics (e.g., granular business statistics).

Effects perceived: More up-to-date and granular statistics would allow policy makers to better assess the impacts on competition.

- In the area of trade, the digital dimension of service trade or intra-firm trade for goods have been highlighted as having significant gaps in official statistics. International data flows between different countries or regional blocks are also difficult to assess. E-commerce is also becoming more and more important and various trends among consumer habits or unfair practices are hard to measure and to address without accurate statistics on cross-border e-commerce in particular.

Effects perceived: Insufficient evidence to back trade agreements can be a major issue. There is a substantial risk of committing errors or deciding on wrong priorities. When it comes to e-commerce, surveys on ICT use are not very effective as respondents rarely remember many details about their habits. It has been shown that these traditional means of gathering data are often ineffective, and without clear and accurate data on consumer habits and cross-border sales, it is very difficult to identify harmful practices or to better target consumer protection measures.

- In the area of defence policies, certain gaps identified were data on the supply and demand side of the defence market (defence and security companies).

Effects perceived: Statistics that are currently available do not always help creating a rigorous basis for the decision-making in this area, which forces users to employ several workarounds to inform their work.

- In the area of transport, data on passenger mobility, especially by car and public transport is one of the gaps highlighted. These data are collected on a voluntary basis but there are many gaps (almost half of the countries are missing). Data for missing countries need to be estimated using different sources.

Effects perceived: In order to have complete data at EU level, this requires filling in the gaps with data from other sources or studies or estimates, which means that the indicators are not fully comparable. Collecting them through studies or surveys requires more financial resources.

Benefits of new or improved statistics that make use of private data sources to close the information gaps

Among the most identified benefits of new or improved statistics is, of course, better policymaking. Generally speaking, statistical users report that having access to better evidence would always improve the quality of the analysis and the quality of the policies that are designed.

Access to more granular privately held data could also enhance the knowledge of policy makers about the trends that they observe and assess and as a result, they can better target their investments and policies.

Moreover, statistical users argue that the main benefit for them would be to use real time data to inform policymakers about "real issues" during a crisis. More generally, collecting data from new sources would be beneficial to suggest when policy should intervene and correct market failures, but also to know when there is the risk of government failure. There will be clear societal benefits if they could understand the extent of certain policy issues or to know that emerging technologies are not causing harm.

Several instances have also emerged where data coming from the private sector are seen as beneficial for statistical users and policy makers:

- Private data would be useful to monitor income. For this, various information from the banking sector could be of interest, such as transaction data coming from credit cards or debit cards. This kind of data could provide useful insights on consumption and therefore anticipate some trends on incomes as well.
- Data on digital platform employment can come directly from platforms themselves, if B2G4S data sharing is possible. This is an aspect that sits high on the political agenda of the European Commission as part of the Digital Economy and Digital Transition topics.
- Data on general trends about the location of public transport stops and departures and arrivals is extremely useful for transport policy. A variety of public and private actors could provide access to those data. Population daytime grids would be very useful to understand where people spend most of their time during the day. That will inform flows, exposures to risks to pollution, etc. and will also give a lot of information on transport needs, transport flows. Any kind of data that mobile phones could provide that give the length of trips, the number of trips or the origin and destination for trips would be a huge benefit for understanding where more transport investments are needed. Data on passengers using cars could also be improved with the help of GPS data from MNOs and could improve policymaking in the area of transport. This kind of data from the private sector would help policy makers understand the impact of the green transition, which is extremely relevant for society.
- Getting data on the amount of energy consumed, which could come from smart meters would be very useful in order to know how much electricity is fed back into the grid coming from solar panels, for example. Again, this would have a huge positive impact on the green transition measures. The coupling of energy data with census data would be most helpful to better understand matters including the structure of families, the energy performance of buildings, and patterns of energy consumption. This would, in turn, be a significant benefit in helping to provide evidence for certain policy priorities including climate mitigation and wider environmental policies. Moreover, having access to such data would be of particular help for cohesion policy.
- Another example is data on the value of the housing or the cost of rent. Right now, Eurostat collects that data in a few different ways, but access to more comprehensive subnational data would be extremely helpful to policy makers. Some private companies have access to a large set of transactions on the housing market and with that information, one can understand those patterns much better.

- Data coming from farms would be extremely beneficial for policymaking in the area of agriculture, as farms are better and better tracked. Inputs on fertiliser, pesticides, water, supply chains and livestock would be game changers for policy in this area. Certain new technologies, such as AI, as well as having an integrated digital reporting mechanism would make this kind of tracking cost-effective and cost-efficient, because on the medium to long term, it would reduce reporting burdens on farms and businesses.
- For tourism statistics, people who travel and who use travelling and tourism services but don't necessarily spend a night in accommodation, are not always visible. Here the possibility appears to use mobile phone data, the possibility to use credit card data, to supplement the gaps. These types of data would help policy makers to recognise more clearly the economic impact, value of tourism and the flows of people between the countries for tourism and travelling. Having access to privately held data would help to have more extensive and timely knowledge of what is happening in travelling, tourism. That would help policy makers to recognise if there are some sudden trends or if there have been shocks, to recognise what are the impacts of those shocks. The data would also help them to understand better what specific support might be needed for tourism because this is a very regional phenomenon, and this would benefit the tourism ecosystem and Europe as a whole.
- Statistics produced with the help of data from labour platforms provide detailed information on the existing vacancies, which is very timely and very useful. In the pandemic crisis, seeing a close and near real-time observation of labour markets was, of course, the key to understanding how economic growth is, how employment, and unemployment is developing.
- The main benefits of business metrics coming from the private sector would be the added credibility, increased timeliness and the ability to account for short-term and structural changes when assessing productivity and competitiveness, for example, when necessary to analyse and prepare for very big structural changes (i.e., exogenous shocks such as Covid-19). More comprehensive and harmonised databases on goods, trade and investments (covering the EU but also other countries) could create a “complete picture” for trade negotiations.

ANNEX 8: SME TEST

The initiative is considered as relevant for SMEs.

Step 1/4: Identification of affected businesses

A distinction is needed between businesses as data users, data survey respondents and data holders.

The more timely and detailed statistics that the initiative will bring about, will benefit all businesses, including SMEs, in their role as data users.

The use of by-product of digital services for European statistics will minimize the need for data collection via business surveys. As such, the initiative will benefit all businesses through a lower response burden, but since SMEs make up the majority of respondents in business surveys, this category of businesses will benefit more than other categories.

Still, businesses that are also data holders, will also be negatively affected by the initiative because of requests for access to their data. However, those businesses will almost exclusively be large businesses. This is because of the market concentration of digital services (mobile network operators, banks (financial transaction service providers), smart meter operators, web portals, etc.).

Step 2/4: Consultation of SME Stakeholders

The consultation covered SMEs with respondents in the survey being asked to assess the overall impact of each policy option on the operation and competitiveness of SMEs and micro-enterprises in particular. All stakeholders preferred policy option 1 to policy option 2, but SMEs expressed concerns in case they would receive requests for access to their data, since this would likely increase their costs, given their limited resources. However, SMEs would welcome burden reduction if data sharing by large enterprises would reduce the need for surveys.

Moreover, the use of big data sources will improve the timeliness of the statistical production, which could increase the competitiveness of SMEs.

Stakeholders further discussed the impact of the policy options on the operation and competitiveness of SMEs and micro-enterprises in the online stakeholder workshop. While participants did not generally provide additional evidence, one NSI suggested that, in the long term, SMEs could expect a reduction of the burden as a result of receiving fewer requests to complete surveys or other administrative forms (if mandatory access to privately-held data is introduced).

The membership of the Expert Group on facilitating the use of new data sources for

official statistics was chosen from different stakeholder groups, including SMEs. The unanimous conclusions of this Expert Group are an important input to the initiative to revise Regulation (EC) No 223/2009 and have been used in this impact assessment.

Step 3/4: Assessment of the impact on SMEs

The baseline option of PO0 does not directly change the burden on SMEs. However, there is also no lowering of the burden on SMEs through alternative ways of data collection. As such, this policy option is neutral in terms of both burden and benefits.

Under PO1 the distribution among businesses of costs and benefits would be uneven. Data holders would incur costs, but since SMEs will only rarely be holders of data that will be in scope for requests for access, this will normally not affect the SMEs. Moreover, PO1 includes a blanket exemption of mandatory data sharing for small and micro enterprises. Turning to the benefits of PO1, the reduction of burden will affect all sizes of businesses, including SMEs. The estimations show that the savings due to decrease of samples are more than 10 time higher than the additional burden on enterprises due to new data demands, and SMEs would particularly benefit from PO1, since they make up the majority of respondents in business surveys.

Under PO2, the distribution among businesses of costs and benefits would be uneven, as was the case for PO1, and as for PO1 SMEs will only rarely be subjected to data requests. On the benefit side, PO2 is also comparable to PO1, where SMEs would particularly benefit, since they make up the majority of respondents in business surveys. However, the response reduction is estimated to be considerably less for PO2 compared to PO1.

Still, under PO2 more enterprises will be affected in terms of burden, compared to PO1, by requests to provide data in cases of urgent user demands.

Step 4/4: Minimising negative impacts on SMEs

Some measures have been considered to mitigate the impacts on SMEs. These are:

- Micro and small enterprises deserve extra safeguards that they will not incur costs due to mandatory data sharing. Therefore, in the preferred option a threshold in terms of the size of businesses will apply to mandatory data sharing. For surveys the decision on thresholds for businesses as data providers is taken in the context of the AWP on the basis of its effect on the contents and quality of the statistics concerned and the associated public benefits. However, for holders of private data as by-products of digital services, a blanket exemption for micro and small enterprises is included in the preferred option, PO1. This is justified, since it provides the strongest safeguard conceivable.