

Business-to-Government Data Sharing for Public Interests in the European Union: Results of a Public Consultation

Iryna Susha¹, Jakob Schiele¹ and Koen Frenken¹

¹ Utrecht University, Copernicus Institute of Sustainable Development, Princetonlaan 8a, 3584
CB Utrecht, The Netherlands
i.susha@uu.nl

Abstract. Lately governments and companies began experimenting with voluntary data sharing of business data for addressing public problems (so-called Data Collaboratives). This early practice revealed a number of challenges impeding business-to-government (B2G) data sharing and thus limiting the potential of data to provide answers and guide policies and action. One of the key challenges is the lack of a clear regulatory framework for B2G data sharing. To tackle this issue, the European Commission is taking regulatory action and preparing the Data Act which aims to spell out the rules and conditions for B2G data sharing for public interest. These developments, however, are met with resistance. While there is a strong push from the public sector for more private sector data, the private sector is less enthusiastic about the prospective mandatory B2G data sharing. In our study we zoom in on this issue in more detail and pose the following research question: How do public and private sector actors in the European Union view the prospect of mandatory B2G data sharing for public interest? To answer this question, we analyze the open dataset of responses to the public consultation of the European Commission. We find statistically significant results of business opposition to regulatory action and to mandating B2G data sharing, particularly among telecom and finance sectors. We also conclude that opposition to mandatory data sharing varies depending on the public interest purpose and is lowest among businesses with regards to emergencies and highest with regard to education, inclusion, and statistics.

Keywords: Data Sharing, Business-to-Government, e-Consultation, European Union

1 Introduction

In recent years we have witnessed the rise of the so-called ‘data for good’ movement. This movement pursues the use of data for societal benefit and its normative orientation is that data should serve public interest. This narrative became widespread especially in the context of grand societal challenges and so-called wicked problems. Data is seen

as a crucial piece of the puzzle in enabling evidence-based decisions (Choi, Gil-Garcia et al. 2021) and data-driven innovation. Much of the data that can potentially be useful for addressing societal challenges is held by the private sector (Noveck, 2015). Unlocking this data from behind the corporate walls is challenging.

Lately governments and companies, often in partnerships with NGOs and research organizations, began experimenting with voluntary data sharing of business data for addressing public problems. These partnerships, termed “data collaboratives” (Susha, Janssen, & Verhulst, 2017), created momentum for collaborative problem-solving of grand challenges, however, this early practice also revealed a number of challenges impeding business-to-government (B2G) data sharing and thus limiting the potential of data to provide answers and guide policies and action. The key problem is that, while businesses are willing to provide data on an occasional basis, they lack incentives to scale up such ‘data philanthropy’ and expect to receive gains in some form from these collaborations (European Commission, 2020). Other barriers concern the lack of clear regulatory framework for B2G data sharing.

To tackle the latter issue, the EU is taking regulatory action and preparing the Data Act which aims to spell out the rules and conditions for B2G data sharing for public interest. The Data Act is the key pillar of the European Strategy for Data and, among other things, aims to create the means for public sector bodies to access and use data held by the private sector that is necessary for specific public interest purposes¹ and mandate such sharing in specific situations.

These developments, however, are met with resistance. While there is a strong push from the public sector for more private sector data, the private sector is less enthusiastic about the prospective mandatory B2G data sharing. The preliminary analysis of the public consultation by the European Commission showed that when it comes to making B2G data sharing mandatory, the preferences of public and private sectors vary. 91% of public authorities consider that regulatory action (EU or national) on B2G data sharing is needed, while this is the opinion of just 38% of business actors (European Commission, 2021). Furthermore, there appears to be disagreement between public and private sector respondents regarding in which key areas B2G data sharing should be made mandatory (Ibid.). We find these initial results intriguing and in our study we zoom in on the following research question: *How do public and private sector actors in the EU view the prospect of mandatory B2G data sharing for public interest?* To answer this question, we analyze the open dataset of responses to the public consultation of the EC using regression analysis.

The phenomenon of mandatory B2G data sharing for public interest is a recent one and research on this topic is limited to a handful of grey literature publications. Current academic knowledge about mandatory B2G data sharing is limited to situations of compliance, such as financial reporting (Troshani, Janssen et al. 2018), and the literature on B2G data sharing for public interest is limited to the situations of voluntary data sharing (Susha, Rukanova et al. 2019, Rukanova, Tan et al. 2020). Thus, our research aims to fill a literature void on mandatory B2G data sharing for public interest.

¹ Data Act Press Release, <https://digital-strategy.ec.europa.eu/en/policies/data-act>

The paper is structured as follows: in section 2 a review of relevant literature on B2G data sharing is presented, followed by the description of our research method in section 3. In section 4 we present the results of our analysis which are discussed in section 5 in view of implications and significance for research. We conclude the paper with key points in section 6.

2 Modalities of B2G data sharing for public interest

To date B2G data sharing for public interest has been mainly realized by means of voluntary data sharing arrangements. In research two such approaches to voluntary data sharing are discussed: data donorship, also known as data philanthropy (George et al., 2020, 2022), and data collaboratives (Susha et al., 2017; Klievink et al., 2019), seen as more collaborative initiatives to share data based on mutual interests of government and business (Micheli, 2022). Research on data philanthropy views this phenomenon as a subtype of corporate philanthropy and focuses on understanding the benefits to donor firms themselves (Awasthi & George, 2019; George et al., 2022). For instance, George et al. (2020) argue that data philanthropy is beneficial to firms and can improve organizational effectiveness but it requires high level of control of the complementary assets (e.g. data expertise). This thesis on the relations between private interest and public value is elaborated in critical literature on digital platforms which questions the common good motives of platform companies (Van Dijck et al., 2018).

Research on data collaboratives views private sector data sharing from a more collective standpoint and strives to identify outcomes at both organizational and societal level (Susha & Gil-Garcia, 2019; Susha, Rukanova et al., 2019). Literature on data collaboratives makes a point that the interests and positions of public and private sector actors differ and might even clash in public-private collaborations (Klievink et al., 2018). While the public sector is driven by the realization of public sector values (Banister, Connolly 2014), the rationale of the private sector remains largely economic and market-driven. Although in principle altruistic motives underlie data partnerships, research shows that companies can often seek to achieve indirect benefits as a bonus of collaborations (Micheli 2022; Susha, Rukanova et al. 2019). Therefore, some arguments have been put forward that B2G data sharing needs to be mutually beneficial and aim for a win-win arrangement (European Commission, 2020). Rukanova et al. (2020) for instance demonstrated how win-wins can be realized by leveraging government leadership and aligning the needs and interests of the different parties. In a similar vein, Susha et al. (2019) conceptualized three different partnerships models wherein self-interest of actors and the societal interest at stake can configure differently. Klievink et al. (2018) demonstrated the success of a data partnership through the mechanisms of collaborative governance and the critical role of trust in public-private relations. Broader literature on inter-organizational collaboration supports this thesis, for instances, as theorized in the work of Porter & Kramer (2011) on shared value or in the framework of collaborative value creation in cross-sector partnerships (Austin & Seitani, 2012).

There are other less explored modalities of B2G data sharing, such as public procurement of data and tender obligations mandating the supplier of a given service to share their data with a government authority (Micheli, 2022). Research on the procurement of private sector data is, to the best of our knowledge, hard to come across. Regarding the latter, the city of Barcelona is becoming a frontrunner in this direction, taking action to introduce the so-called data sovereignty clauses in procurement contracts which would enable the city to access data ‘about the city’ from private sector providers (Monge et al., 2022). Mandatory B2G data sharing for public interest has not received much attention in the literature yet. However, there has emerged a broader narrative on data sovereignty (Hummel et al., 2021) and the need to give back control of data to citizens. Issues of power dynamics between public and private actors (Micheli, 2022) and in relation to the citizens (Mercille, 2021) have also been brought to the surface. In a qualitative study of 12 European cities Micheli (2022) found that public bodies often lack the means to set the terms of how data is shared, thus it is not assured the information they get will be useful to their public interest purposes. In this respect the public actors have the role of passive ‘recipients’ of data (Ibid.) and have to deal with a take-it-or-leave-it situation. This dependence on the private sector for data produces emerging negative effects, such as for instance it introduces inequalities between public bodies that were able to access business data and innovate and those that were not (Ibid.).

Having said that, we know little about the interplay between the aforesaid modalities of data sharing, how the actors settle for a particular data sharing mode, and which approach should be favored in any given situation. Vigorito (2022), for instance, makes a point that there is potential for combining voluntary and mandatory modes into a ‘hybrid’ approach, but how this should be done remains for further exploration. Furthermore, existing research paints a rich picture of drivers and barriers for data sharing between government and business in general (Sayogo & Pardo, 2013; Klievink et al., 2016; Susha, Gronlund & Van Tulder, 2019). Yet, a more nuanced and actor-specific investigation of how public versus private actors experience and view B2G data sharing is needed. The research by Micheli (2022) exploring the experiences and views of local administrations on B2G data sharing is a step in this direction. Our study aims to add to this knowledge by investigating the views of public and private actors regarding the prospect of mandatory B2G data sharing for public interest in the EU.

3 Methodology

To analyze how public and private actors in the EU view the prospect of mandatory B2G data sharing for public interest, we look at survey data collected by the European Commission in a public consultation on the Data Act regarding B2G data sharing. In an online questionnaire, the Commission asked respondents to answer 158 questions and to provide some general information about the respondents. This resulted in an extensive database, including quantitative and qualitative data on 449 respondents,

which is publicly accessible on the Commission website². For each respondent, we extracted their answers to eight questions about B2G data sharing, as well as some background information on the respondents (type, size, country, sector).

The main relevant question on B2G data sharing for public interest in the EU is question 3 of the survey (“Should the EU take additional action so that public sector bodies can access and re-use private sector data, when this data is needed for them to carry out their tasks in the public interest purpose?”). The answer to this question best captures an actor’s general attitude towards regulating B2G data sharing. We only consider respondents who are either favorable (“EU level action is needed”) or opposing (“No action is needed”), discarding 15 respondents who express an ambivalent position (“Action at the Member State level only is needed”) and 113 respondents who did not provide any response. As our main dependent variable, we created a dummy variable *General* which takes on 1 for respondents who generally oppose regulatory action and 0 for respondents who are generally in favor of regulatory action.

We further consider questions 15-21, which have the same structure as question 3, but are specific to a particular public interest in which data sharing should be mandated (“In which of the following areas do you think that, for specific use-cases with a clear public interest, B2G data sharing should be compulsory, with appropriate safeguards?”). Here, possible responses are binary (“Yes, it should be compulsory” and “No, it should not be compulsory”), again discarding respondents who did not leave a response. We created, similar to the dependent variable *General*, seven dummy variables as dependent variables, which take on 1 for respondents indicating opposition to compulsory data sharing for a specific public interest, and 0 for respondents indicating no opposition to compulsory data sharing for that interest. The seven dummy variables related to the seven public interests are:

<i>Emergency:</i>	“Data (e.g. mobility data from Telecom operators, loss data from insurance companies) for emergencies and crisis management, prevention and resilience”;
<i>Statistics:</i>	“Data (e.g. price data from supermarkets) for official statistics”;
<i>Environment:</i>	“Data (e.g. emissions data from manufacturing plants) for protecting the environment”;
<i>Health:</i>	“Data (e.g. fuel consumption data from transport operators) for a healthier society”;
<i>Education:</i>	“Data for better public education services”;
<i>Inclusion:</i>	“Data (e.g. employment data from companies) for a socially inclusive society”;
<i>Policy:</i>	“Data for evidence-based public service delivery and policy-making”.

The background information on the respondents include actor type (*Type*) which distinguishes between (A) “Business Association” (n=122), (B) “Company/Business organisation” (n=105), (C) “Public authority” (n=100), (D) “EU Citizen” (n=56), (E) “NGO (Non-governmental organisation)” (n=21), (F) “Academic/Research Institution” (n=17), (G) “Consumer Organisation” (n=6), (H) “Non-EU Citizen” (n=2), (I) “Trade Union” (n=2) and (J) “Other” (n=18). As we were mainly interested in differences between government and business actors, we created the dummy variables *Business*

² https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/13045-Data-Act-&-amended-rules-on-the-legal-protection-of-databases/public-consultation_en

(grouping together (A) and (B)) and *Public* (equal to (C)). For the large number of citizens, we created a third dummy *Citizen* (grouping (D) and (H)). All other actor types are used as reference category in the regression analysis.

Organization size (*Size*) is an ordinal variable including “Large (250 or more)” (n=171), “Medium (< 250 employees)” (n=77), “Small (< 50 employees)” (n=74) and “Micro (< 10 employees)” (n=69), with 58 NAs. As these 58 non-responses coincided with the user type citizen, we decided to group them together with Micro, as a citizen can be regarded as representing a single person. We use as dummy variables *Large*, *Medium* and *Small*, with *Micro* as reference category.

For the respondent’s country (*Country*), a total of 32 countries were mentioned, including some non-EU countries. However, many of these had very low numbers of observations. Therefore, we created dummy variables only for countries with 10 or more respondents to question 3 (on which the main dependent variable *General* is based.) This leads to ten country dummies: *Belgium* (n=44), *Finland* (n=13), *France* (n=35), *Germany* (n=60), *Hungary* (n=11), *Italy* (n=18), *Lithuania* (n=13), *Netherlands* (n=16), *Poland* (n=10) and *Spain* (n=49). All the other countries serve as the reference category.

Similarly, for business sector (*Sector*) we created dummy variables only for sectors with 10 or more respondents. These are: “Agriculture, forestry and fishing” (*Agri*, n=10), “Automotive, including suppliers, manufacturing, retail, service and maintenance and related after-market services” (*Auto*, n=13), “Finance, insurance and re-insurance (other than motor insurance)” (*Finance*, n=19), “IT” (*IT*, n=42), “Legal advice; market research” (*Legal*, n=11), “Media, publishing, broadcasting and related services including advertising” (*Media*, n=10), and “Telecommunications, including suppliers” (*Telecom*, n=10). All other sectors serve as a reference category.

Below, we present descriptive statistics, as well as the results from eight binary regression models, all using the same list of dummies as independent variables (*Type*, *Size*, *Country* & *Sector*), but different dependent variables (*General*, *Emergency*, *Statistics*, *Environment*, *Health*, *Education*, *Inclusion* and *Policy*). Variance Inflation Factors (VIF) were calculated for all regression models, indicating that there is no multicollinearity (values all well below 10). The analysis was done in Rstudio version 4.0.3 (R Core Team, 2020), where for data manipulation the package *dplyr* (Wickham, François, Henry & Müller, 2021) was used next to the base version and VIF using the package *car* (Fox & Weisberg, 2019).

4 Results

Table 1 shows the descriptive statistics for all variables. Looking first at all respondents, we see that 29 percent opposes regulation in general, while this percentage tends to be higher for specific public interests. In particular, there is strong opposition against mandatory data sharing for inclusion, which involves sharing of employment data. Apparently, many actors do not feel that mandatory sharing of such data is justified.

Table 1. Percentage of respondents opposing B2G data sharing regulation								
	% opposed							
variable	<i>General</i>	<i>Emerg.</i>	<i>Stat.</i>	<i>Env.</i>	<i>Health</i>	<i>Educ.</i>	<i>Inclus.</i>	<i>Policy</i>
All	29.0	30.2	37.7	26.7	34.4	35.4	47.1	37.0
Business	54.4	54.1	75.2	55.7	58.7	66.3	80.0	68.1
Citizen	10.0	18.0	25.5	13.2	35.3	33.3	42.2	32.6
Public	1.3	5.8	3.3	1.2	6.3	6.6	9.5	4.8
Other	13.9	16.2	20.5	7.8	13.8	9.3	18.1	12.1
Large	31.1	26.1	30.6	21.9	25.9	30.4	40.8	32.1
Medium	27.5	24.0	31.4	22.5	29.2	23.8	36.2	27.1
Small	39.1	40.4	52.3	37.2	42.1	44.1	59.5	47.6
Other	21.7	33.7	42.8	29.7	43.6	42.7	54.1	43.6
Belgium	50.0	62.8	74.3	55.3	56.8	69.0	80.0	70.3
Finland	23.1	7.1	46.7	30.8	30.8	45.5	57.1	33.3
France	25.7	37.0	40.9	32.0	41.7	38.9	57.9	44.0
Germany	38.3	33.0	41.2	30.2	37.3	38.1	53.5	36.0
Hungary	18.2	8.3	7.7	0.0	7.7	16.7	18.2	25.0
Italy	16.7	17.7	52.9	11.1	17.7	26.7	18.8	28.6
Lithuania	23.1	17.7	5.9	10.5	31.3	22.2	22.2	17.7
Netherlands	31.3	58.3	58.3	38.5	41.7	45.5	58.3	46.2
Poland	20.0	33.3	20.0	22.2	44.4	22.2	22.2	28.6
Spain	8.2	7.3	7.7	5.9	12.2	14.6	17.4	16.3
Other	32.6	32.6	47.1	34.0	44.6	43.1	59.1	41.6
Agriculture	30.0	33.3	30.0	12.5	33.3	10.0	50.0	30.0
Automotive	30.8	44.4	91.7	55.6	46.7	63.6	92.9	75.0
Finance	73.7	40.0	66.7	25.0	25.0	33.3	54.6	72.7
IT	35.7	53.7	74.3	50.0	61.1	64.7	68.6	68.6
Legal	18.2	12.5	50.0	20.0	55.6	37.5	66.7	33.3
Media	60.0	66.7	75.0	50.0	66.7	55.6	71.4	75.0
Telecom	80.0	80.0	83.3	80.0	66.7	66.7	83.3	80.0
Other	19.9	20.2	24.1	18.5	25.5	27.1	35.7	24.6

Table 2: Regression results									
Regression model		Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
DVs: opposed to sharing in		general	emergencies	statistics	environment	health	education	inclusion	policy
Independent Variables									
Type	Business	2.068 ***	1.855 ***	3.704 ***	3.035 ***	2.586 ***	3.362 ***	3.494 ***	3.066 ***
	Citizen	-0.320	-0.204	-0.207	0.226	0.907	1.516 *	0.938	1.028
	Public	-2.795 **	-0.537	-0.524	-1.211	-0.290	0.233	0.066	-0.406
Size	Large	0.098	-0.774	-1.456 **	-0.937 *	-0.858	-0.419	-0.801	-0.601
	Medium	-0.412	-1.340 **	-2.523 ***	-1.366 **	-1.344 **	-1.413 **	-1.942 ***	-1.425 **
	Small	0.027	-0.711	-1.800 **	-0.990	-0.799	-1.073	-1.024	-0.975
Country	Belgium	-0.168	0.978 *	0.771	0.243	-0.000	0.798	0.531	0.701
	Finland	-0.164	-1.550	0.488	-0.004	-0.344	0.890	0.459	
	France	-1.705 ***	-0.320	-1.681 **	-0.844	-0.798	-0.662	-0.647	-0.706
	Germany	-0.015	0.530	0.122	-0.137	-0.198	0.384	0.200	-0.105
	Hungary	-1.017	-1.726	-2.113 *	-16.292	-2.071 *	-0.848	-1.720 *	0.018
	Italy	-1.551 *	-0.672 *	0.605	-1.726 *	-1.637 **	-0.700	-0.765	-0.389
	Lithuania	1.073	0.306	-2.139	-0.139	0.663	0.446	-0.094	0.360
	Netherlands	-1.028	1.141	0.425	0.070	-0.080	0.643	0.120	0.395
	Poland	-0.287		-0.046					
	Spain	-1.193	-1.129 *	-1.885 **	-1.224	-1.112 *	-0.461	-0.809	-0.242
Sector	Agriculture	1.206		1.034			-1.878		0.835
	Automotive	-0.701	-0.176	2.159 *	0.238	-0.379	0.067	1.568	1.017
	Finance	2.118 ***	0.083					-0.504	1.480 *
	IT	-0.057	1.028	2.418 ***	0.809 *	0.853 *	1.180 **	0.647	1.463 ***
	Legal	0.264		2.672 ***	0.459				
	Media	0.832							
	Telecom	2.167 **	2.168 **						
AIC	294.69	303.14	235.42	255.48	305.91	260.72	269.71	277.09	
Observations	321	308	297	300	288	257	272	281	

Significance codes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 3: Odds ratios									
Regression model		Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
DVs: opposed to sharing in		general	emergencies	statistics	environment	health	education	inclusion	policy
Independent Variables									
Type	Business	7.90	6.39	40.64	20.81	13.28	28.87	32.93	21.46
	Citizen						4.55		
	Public	0.06							
Size	Large			0.23	0.39				
	Medium		0.26	0.08	0.25	0.26	0.24	0.14	0.24
	Small			0.16					
Country	Belgium		2.66						
	Germany								
	Spain		0.32	0.15		0.32			
	Finland								
	France	0.18		0.18				0.17	
	Hungary			0.12					
	Italy	0.21	0.51		0.17	0.19			
	Lithuania								
	Netherlands								
	Poland								
Sector	Agriculture			8.66					
	Automotive								
	Finance	8.31							4.39
	IT			11.23	2.24	2.34	3.25		4.32
	Legal			14.47					
	Media								
	Telecom	8.73	8.74						

Turning to specific categories of respondents, it is clear that the main differences in opinion lie between the *Type* of actors. The majority of *Business* actors oppose B2G data sharing regulation, while the large majority of *Citizen* and *Public* actors are in favor. This pattern is repeated when looking at specific interests, where the share of opposing businesses is generally even larger. In relative terms, businesses seem most open to data sharing for emergency and environmental purposes.

Regarding *Size*, we do not discern clear patterns. *Country* responses are generally also rather close, with only *Belgium* and *Spain* as outliers. Sectoral differences are more pronounced with respondents from data-intensive sectors like *Telecom*, *Finance* and *Media* mostly opposing. Respondents from *Legal* services are most strongly in favor, which may reflect that new regulations provide them with more business.

While the descriptive statistics indicate some patterns, one cannot conclude from percentages alone that some types of respondents would oppose B2G data sharing regulation more than others. To reach more conclusive answers, we run logistic regressions, including all variables. In this way, we can assess whether certain types of respondents significantly more often oppose, while controlling for all other factors.

Table 2 shows the results of the binary logit regression models and Table 3 shows the corresponding odds ratios for the significant effects ($p < 0.1$). Note that we only include dummy variables in the regressions with 10 or more observations, which implies that in Models 2-8 some coefficients are missing.

Model 1 with *General* as dependent variable shows the results for the answers to the general question regarding regulation. We observe a highly significant positive effect of *Business* (coef= 2.068, $p < 0.01$). In terms of the odds ratio, business actors are 7.9 times more likely to oppose regulatory action on B2G data sharing. Inversely, *Public* has a negative effect on opposition (coef= -2.795, $p < 0.05$), with public authorities being 16.7 times less likely to oppose B2G data sharing regulation. This confirms the strongly diverging interests and positions held by business actors and public authorities regarding data sharing (Klievink et al., 2018).

Size did not show any significant effect in *Model 1*, so there does not seem to be a discernible difference in attitudes towards general data sharing related to organization size. Looking at *Country* effects, *France* stands out as having a highly significant negative effect on opposition (coef= -1.705, $p < 0.01$) and Italy also has a negative effect, albeit less significant (coef= -1.551, $p < 0.1$). Actors residing in France are 5.6 times less likely and those in *Italy* 4.8 times less likely to oppose data sharing regulation. This fits with the notion of ‘varieties of capitalism’ which characterizes these countries as a Mediterranean variety, marked by “histories of extensive state intervention” (Hall & Soskice, 2001, p. 21). Lastly, respondents from *Finance* (coef= 2.118, $p < 0.001$) and *Telecom* (coef= 2.167, $p < 0.05$) tend to oppose regulation. Finance actors are 8.3 times more likely to oppose data sharing regulation and telecom actors 8.7 times more likely to oppose. In summary, the strongest opposition to regulatory action on B2G data sharing in general can be found in business actors, and specifically, in the finance & insurance and telecommunications sectors.

Looking at the results for specific public interests in Model 2-8, it becomes clear that business actors’ strong opposition to compulsory data sharing is consistent across the

different subject areas, although it is the least strong in *Emergencies*, and strongest for *Statistics*, followed by *Inclusion* and *Education*.

The coefficients for *Size* show that across the subject areas, *Medium* sized organizations tend to be least opposing regulation. In some areas, *France* and *Italy* again appear as having a negative effect on opposition, but overall a clear country pattern is missing. Finally, we see that the *IT* sector seems to oppose mandatory data sharing for most public interests. This may reflect the data-intensive nature of the IT industry, with the data as a primary resource for their operations and innovations.

5 Discussion

Our research provides convincing **empirical evidence of the divide** in views of government and business on the prospect of mandatory B2G data sharing for public interest in the EU. Our analysis of the results of the public consultation on the Data Act shows that the majority of business actors oppose regulatory action on B2G data sharing for public interest.

The **stronghold of opposition** to mandatory B2G data sharing has been identified to be in the telecom and finance sectors. Interestingly, these sectors already have a track record of ‘success stories’ in voluntary data sharing collaborations with the government. Telecom data has been named as one of the most sought-after types of data by government authorities, according to research by Micheli (2022). At the same time, telecom data has high value as a proprietary resource that contains insights about business processes and customer preferences (Taylor, 2016). Nonetheless, telecom companies have been rather active lately in the practice of voluntary data sharing through data philanthropy or data collaborative initiatives. In 2017 the GSMA established the Data for Social Good initiative in which 16 world leading mobile network operators joined to leverage their big data capabilities in order to address humanitarian crises. During the Covid-19 pandemic the GSMA was approached by the European Commission with a request to share anonymized and aggregate mobile positioning data in order to study and respond to the pandemic. As a result, 17 mobile operators shared the data with the EC covering 22 EU member states and Norway which was seen as an initiative of “unprecedented nature” (Vespe et al., 2021). Other ‘success stories’ of voluntary data sharing of telecom data in the times of the pandemic can be found in a special issue of *Data & Policy* (Benjamins, Vos, & Verhulst, 2022). The guest editors of the special issue (two of whom are representatives of the mobile industry) reflected on these practices and formulated two key challenges: financing to ensure long-term supply of data and building capacities and digital skills among government organizations (Ibid.). An interesting case is described by Agren et al. (2021) when an initially non-profit offering of a data analytics product (Crowd Insights) by Telia to the Public Health Agency of Sweden was turned into a commercial contract as the pandemic continued. There is evidence of the same strategy in the research by Micheli (2021) where businesses approach government organizations with a free offering which they afterwards transform into a

commercial product. Telecom data thus presents an instance where high commercial value and high potential for societal impact collide.

Our second observation is that **resistance to data sharing varies depending on the public interest purpose**. Our findings show that sharing data for emergencies was met with least opposition from the business side, whereas such public interest purposes, as statistics, inclusion, and education received most resistance. Overall, companies around the world have been engaging in emergency-related activities as part of corporate social responsibility. Research shows that such involvement is based on both ethical/moral motives and instrumental motives (business profitability, continuity) (Johnson et al., 2011). At the same time, the strongest ‘pull’ for data on the side of the public sector is for data for statistical purposes, emergencies, and environment (European Commission, 2021). Our findings point to a **mismatch in terms of demand in the public sector and the willingness to disclose** on the business sector side. In the case of official statistics, so far statistics agencies have experimented with accessing private sector data via voluntary data sharing arrangements with companies; a variety of data sharing models to this end are proposed (Klein et al., 2016; Klein & Verhulst, 2017). According to OECD, challenges in these B2G data sharing partnerships for statistics include competitive risks, lack of incentives and turning it into a viable business model, and reputational and ethical risks (Klein et al., 2016). Next to statistics, our findings also point to low levels of enthusiasm among business about mandatory data sharing for education purposes. At the same time, there are rising concerns about the datafication of education and the reliance on private technology providers. Olazp et al. (2022) describe the process of “digital colonization” by digital platforms of the education sector and warn about the changing power dynamics between the incumbent service providers and the platforms where the latter control the data and become unique providers of critical, data-driven value.

Our research demonstrated that when it comes to B2G data sharing private interest and public value are in conflict. Van Dijk et al. (2018) argue that government has a role to play in safeguarding the public value by acting as a regulator, enforcer, or stimulant of public value. At the time of writing, on 23 February 2022, the EC published the text of the proposal for the Data Act, having considered input from stakeholder consultations, including the said public consultation. The proposal mandates the sharing of business data with government in situations of “exceptional need” thus adopting a narrow approach to public interest and favoring more the position of business. The proposal is yet to go through the legislative process of adoption. However, our research shows that the future of B2G data sharing in the EU will be situated in a highly rivalrous landscape where the interests and views of the public and private sectors are extremely divided.

In our future research we aim to enrich the quantitative analysis presented in this paper with insights from qualitative analysis of submissions which accompanied the public consultation responses. This will enable us to shed more light on the underlying arguments of actors in support of or in opposition to mandatory B2G data sharing in the EU.

6 Conclusions

In our research we focused on the state of B2G data sharing in the EU, and namely on the forthcoming EU Data Act and the prospect of mandatory B2G data sharing for public interest. Our analysis focused on the responses to the public consultation with the aim to answer the question: *How do public and private sector actors in the EU view the prospect of mandatory B2G data sharing for public interest?* We produced descriptive statistics of the results, together with eight binary regression models. We found statistically significant results of business opposition to regulatory action and to mandating B2G data sharing, particularly among telecom and finance sectors. We also concluded that opposition to mandatory data sharing varies depending on the public interest purpose and is lowest among businesses with regards to emergencies and highest with regard to education, inclusion, and statistics. We found that there is a mismatch in terms of the needs articulated by the public sector respondents and the willingness to disclose on the side of the private sector. Our findings paint a picture of divergent positions on B2G data sharing for public interest in the EU and call on future research to take a more critical stance on data and the role of the public sector in the data economy.

Studies of public-private relations in digital government research have overly focused on collaboration and finding synergies thereby following a more ‘optimistic’ curve. Our study exposes the clashing interests and the power struggle in positions of business and government on data and public interest. Our study also adds to a handful of quantitative studies on B2G data sharing (e.g. George et al., 2022). Research on B2G data sharing is heavily based on case studies, and even more so, on case studies of successful collaborations offering anecdotal evidence of mutually beneficial data sharing arrangements. We call on the research community to give due attention to the conflictual nature of information sharing and scrutinize the tensions and even ‘failures’ in business-government collaborations.

The limitations of our study are in part related to the sample. Whereas the overall sample size is reasonably good, some sectors and countries had to be excluded from the regression analysis due to clearly insufficient observations ($n < 10$). Moreover, some remaining dummies still had a quite low number of observations ($n < 20$). This could be ameliorated in future research by using a larger dataset. Nevertheless, our main conclusions remain robust, as they follow from strongly significant effects based on large numbers of observations. In our research we did not discuss in more depth the citizen perspective which we call on future studies to investigate. Citizens have a stake in the B2G data sharing as data subjects and it is important to compare their views against those of the government and business.

References

1. Choi, Y., Gil-Garcia, J., Burke, G. B., Costello, J., Werthmuller, D., & Aranay, O. (2021). Towards Data-Driven Decision-Making in Government: Identifying Opportunities and Challenges for Data Use and Analytics. In *Proceedings of the 54th Hawaii International Conference on System Sciences*.

2. Noveck, B. (2015). Data collaboratives: sharing public data in private hands for social good. *Forbes*. Accessed from www.forbes.com/sites/bethsimonenoveck/2015/09/24/private-data-sharing-for-public-good/-28dab08b65bb on 10 May 2016.
3. Susha, I., Janssen, M., & Verhulst, S. (2017). Data collaboratives as a new frontier of cross-sector partnerships in the age of open data: Taxonomy development. In *Proceedings of the 50th Hawaii International Conference on System Sciences*.
4. European Commission (2020). Towards a European strategy for B2G data sharing for the public interest. Final report of the high-level expert group on business-to-government data sharing. Accessed from https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=64954 on 17 March 2021.
5. European Commission (2021). Public Consultation on Data Act and Amended Rules on the Legal Protection of Databases. Summary report of the public consultation. Accessed from <https://ec.europa.eu/info/law/better-regulation/> on 17 March 2021.
6. Troshani, I., Janssen, M., Lymer, A., & Parker, L. D. (2018). Digital transformation of business-to-government reporting: An institutional work perspective. *International Journal of Accounting Information Systems*, 31, 17-36.
7. Susha, I., D. Rukanova, B., Ramon Gil-Garcia, J., Hua Tan, Y., & Gasco, M. (2019, June). Identifying mechanisms for achieving voluntary data sharing in cross-sector partnerships for public good. In *Proceedings of the 20th Annual International Conference on Digital Government Research* (pp. 227-236).
8. Rukanova, B., Tan, Y. H., Huiden, R., Ravulakollu, A., Grainger, A., & Heijmann, F. (2020). A framework for voluntary business-government information sharing. *Government Information Quarterly*, 37(4), 101501.
9. George, J. J., Yan, J., & Leidner, D. E. (2020). Data Philanthropy: Corporate Responsibility with Strategic Value?. *Information Systems Management*, 37(3), 186-197.
10. George, J. J., Yan, J., Leidner, D. E., & Awasthi, P. (2022). Does Engaging in Data Philanthropy Impact Business Value?. *Information Systems Management*, 1-15.
11. Klievink, B., Van Der Voort, H., & Veeneman, W. (2018). Creating value through data collaboratives. *Information Polity*, 23(4), 379-397.
12. Micheli, M. (2022). Public bodies' access to private sector data: The perspectives of twelve European local administrations. *First Monday*.
13. Awasthi, P., & George, J. (2019). Harmonizing strategic advantage with social good through Data Philanthropy. In *Proceedings Annual Workshop of the AIS Special Interest Group for ICT in Global Development*.
14. Van Dijck, J., Poell, T., & De Waal, M. (2018). *The platform society: Public values in a connective world*. Oxford University Press.
15. Susha, I., & Gil-Garcia, J. R. (2019, January). A collaborative governance approach to partnerships addressing public problems with private data. In *Proceedings of the 52nd Hawaii International Conference on System Sciences*.
16. Bannister, F., & Connolly, R. (2014). ICT, public values and transformative government: A framework and programme for research. *Government Information Quarterly*, 31(1), 119-128.
17. Porter, M. E., & Kramer, M. R. (2019). Creating shared value. In *Managing sustainable business* (pp. 323-346). Springer, Dordrecht.
18. Austin, J. E., & Seitanidi, M. M. (2012). Collaborative value creation: A review of partnering between nonprofits and businesses: Part I. Value creation spectrum and collaboration stages. *Nonprofit and voluntary sector quarterly*, 41(5), 726-758.

19. Monge, F., Initiative, B. H. C. L., Barns, S., Kattel, R., & Bria, F. A new data deal: the case of Barcelona. UCL Institute for Innovation and Public Purpose.
20. Hummel, P., Braun, M., Tretter, M., & Dabrock, P. (2021). Data sovereignty: A review. *Big Data & Society*, 8(1), 2053951720982012.
21. Mercille, J. (2021). Inclusive smart cities: beyond voluntary corporate data sharing. *Sustainability*, 13(15), 8135.
22. Vigorito, A. (2022). Government Access to Privately-Held Data: Business-to-Government Data Sharing: Voluntary and Mandatory Models. *European Journal of Comparative Law and Governance*, 1(aop), 1-22.
23. Sayogo, D. S., & Pardo, T. A. (2013, June). Understanding smart data disclosure policy success: The case of Green Button. In *Proceedings of the 14th annual international conference on digital government research* (pp. 72-81).
24. Klievink, B., Bharosa, N., & Tan, Y. H. (2016). The collaborative realization of public values and business goals: Governance and infrastructure of public-private information platforms. *Government information quarterly*, 33(1), 67-79.
25. Susha, I., Grönlund, Å., & Van Tulder, R. (2019). Data driven social partnerships: Exploring an emergent trend in search of research challenges and questions. *Government Information Quarterly*, 36(1), 112-128.
26. R Core Team (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
27. Wickham, H., François, R., Henry, L. & Müller, K. (2021). *dplyr: A Grammar of Data Manipulation*. R package version 1.0.7. <https://CRAN.R-project.org/package=dplyr>
28. Fox, J. & Weisberg, S. (2019). *An {R} Companion to Applied Regression, Third Edition*. Thousand Oaks CA: Sage. <https://socialsciences.mcmaster.ca/jfox/Books/Companion/>
29. Hall, P. A., & Soskice, D. (2001). An Introduction to Varieties of Capitalism. *Varieties of Capitalism* (pp. 1–68). <https://doi.org/10.1093/0199247757.003.0001>
30. Taylor, L. (2016). The ethics of big data as a public good: which public? Whose good?. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2083), 20160126.
31. Vespe, M., Iacus, S. M., Santamaria, C., Sermi, F., & Spyrtos, S. (2021). On the use of data from multiple mobile network operators in Europe to fight COVID-19. *Data & Policy*, 3.
32. Benjamins, R., Vos, J., & Verhulst, S. (2022). Mobile Big Data in the fight against COVID-19. *Data & Policy*, 4.
33. Ågren, K., Bjelkmar, P., & Allison, E. (2021). The use of anonymized and aggregated telecom mobility data by a public health agency during the COVID-19 pandemic: Learnings from both the operator and agency perspective. *Data & Policy*, 3.
34. Johnson, B. R., Connolly, E., & Carter, T. S. (2011). Corporate social responsibility: The role of Fortune 100 companies in domestic and international natural disasters. *Corporate Social Responsibility and Environmental Management*, 18(6), 352-369.
35. Robin, N., Klein, T., & Jütting, J. (2016). Public-private partnerships for statistics: Lessons learned, future steps: A focus on the use of non-official data sources for national statistics and public policy. *OECD Development Co-operation Working Papers*. <https://doi.org/10.1787/5jm3nqp1g8wf-en>.
36. Klein, T., & Verhulst, S. (2017). Access to new data sources for statistics: Business models and incentives for the corporate sector. *PARIS21, Discussion Paper No. 10, March 2017*.
37. Ozalp, H., Ozcan, P., Dinckol, D., Zachariadis, M., Gawer, A. (2022). “Digital Colonization” of Highly Regulated Industries: An Analysis of Big Tech Platforms' Entry into Healthcare and Education. Forthcoming in *California Management Review*.

