DNB Working Paper
No. 728 / November 2021

Not all data are created equal - Data sharing and privacy

Michiel Bijlsma, Carin van der Cruijsen and Nicole Jonker

DeNederlandscheBank
EUROSYSTEEM

Not all data are created equal - Data sharing and privacy

Michiel Bijlsma, Carin van der Cruijsen and Nicole Jonker**

* Views expressed are those of the authors and do not necessarily reflect official positions of De Nederlandsche Bank.

Working Paper No. 728

De Nederlandsche Bank NV
P.O. Box 98
1000 AB AMSTERDAM
The Netherlands

November 2021

Not all data are created equal - Data sharing and privacy*

Michiel Bijlsma[a,b], Carin van der Cruijsen[c] and Nicole Jonker[c]

[a] SEO Amsterdam Economics, The Netherlands

[b] Tilburg University, The Netherlands

[c] De Nederlandsche Bank (DNB), The Netherlands

November 2021

Abstract

The COVID-19 pandemic has increased our online presence and unleashed a new discussion on sharing sensitive personal data. Upcoming European legislation will facilitate data sharing in several areas, following the lead of the revised payments directive (PSD2), which enables payments data sharing with third parties. However, little is known about what drives consumers' preferences with different types of data, as preferences may differ according to the type of data, type of usage or type of firm using the data. Using a discrete-choice survey approach among a representative group of Dutch consumers, we find that next to health data, people are hesitant to share their financial data on payments, wealth and pensions, compared to other types of consumer data. Second, consumers are especially cautious about sharing their data when they are not used anonymously. Third, consumers are more hesitant to share their data with BigTechs, webshops and insurers than they are with banks. Fourth, a financial reward can trigger data sharing by consumers. Last, we show that attitudes towards data usage depend on personal characteristics, consumers' digital skills, online behaviour and their trust in the firms using the data.

Keywords: D12; E42; G21; G22; G23

JEL codes: consumer data, data sharing, banks, BigTechs, insurers, webshops, trust, digital skills

* We would like to thank colleagues at DNB for helpful comments on earlier versions of this paper and the questionnaire. We are grateful to Miquelle Marquandt of CentERdata for collecting the data and for her help with the questionnaire. The views expressed in this paper are our own and do not necessarily reflect those of DNB, the ESCB or SEO Amsterdam Economics. All remaining errors are our own.

1. Introduction

Sharing personal data with firms is a central feature of everyday digital life. When people browse the internet, cookies register their website usage, when they use their mobile phones they give up location data and when they pay using debit cards, credit cards or e-wallets, transaction data are recorded. The COVID-19 pandemic has accelerated this development: it has intensified our online life and the amount of data shared via the internet, and ignited a discussion on sharing sensitive health data for use in the fight against the pandemic. Little is known about what drives consumers' agreement with the usage of data by third parties. We add to knowledge on this topic by using a discrete choice survey approach among a representative group of Dutch consumers.

The importance of data has risen sharply in the financial sector and the economy as a whole. With rapidly increasing amounts and varieties of data and the emergence of new technologies enabling large scale data storage and advanced big data analysis, the role of data as a production factor in the economy has grown considerably. Firms and organizations use data to improve existing products and to produce them more efficiently, and also to develop entirely new products. Various studies point to the large social and economic benefits that may arise due to increased data availability and usage by the public and private sector (see e.g. Economic Commission 2020; OECD 2019; McKinsey & Company 2021). As a consequence, the demand for data and hence access to data by the public and the private sector is expected to continue to grow. However, there are also some downsides associated with increased data availability and sharing of people's personal data. For instance, people may be unaware of what the firm may actually do with their data. Furthermore, people's privacy may be at risk if data holding firms excessively use or share these people's data with others, also for purposes for which they did not give their consent. There may also be negative social externalities if sharing of data by one person also leads to disclosure of information from other people who did not give consent to access their data. This not only refers to situations in which other people's data are directly shared, but also to situations in which a sample of individuals from a specific group allows firms to access their data and these firms use their data to derive accurate estimates of the preferences for all people belonging to that group, but who did not disclose their data (see e.g. Choi et al. 2019; Garrat and Van Oordt, 2021).

A relatively recent development is that regulation is actively being developed that allows consumers to decide whether or not they share particular private data with firms to enhance economic growth and welfare from data sharing, while mitigating privacy and other risks. A prime example is the revised Payment Services Directive (PSD2) in the European Union, that regulates access to the payment account for third parties. PSD2 was implemented in 2019 and aims to increase innovation, competition and consumer protection in the European payment market by encouraging current and new service providers to develop and offer new types of services, like account information services. Payment Service Providers (PSPs, often banks) are required to allow licensed third parties access to consumers' (and firms') payment accounts in order to provide payment information or payment transaction services. Consumers have to give their explicit consent to these third parties.

This development can also be seen with respect to other financial and non-financial data, as part of the current impetus towards open finance and to other types of data, such as energy data, telecommunications data or health care data. For example, the EU has formulated a data strategy which aims to create a single market for data.[1] This requires rules and regulation on access to and use of data. In the context of the Digital Finance Strategy, the European Commission announced the intention to adopt a legislative proposal for a new open finance framework by mid-2022.[2] This implies mandating access for third parties to financial customer and business data such as savings or insurance products. In Europe, the UK is at the forefront of open finance with broad adoption by consumers and firms, and the creation of an Open Banking Implementation Entity (OBIE) by the Competition and Market Authority. In Australia, the Consumer Data Right (CDR) was introduced in the banking sector in July 2020 and will be rolled out across other sectors of the economy.[3] In the banking sector, the CDR implies that consumers can share banking data, such as transaction history, interest rates on savings and account balances with third parties. The legislation aims to give Australians the right to access not just financial data, but also their utility and telecoms data. In addition, data sharing of payments data has also already been possible in India since 2016, New Zealand since 2017 and China since 2020 (Swallow et al. 2021).

However, people differ in their willingness to share different classes and types of information and also in the extent to which they trust different types of firms. Little is known about what drives consumers' preferences as to the usage of different types of data, for different types of usage and by different types of firms. We add to knowledge on this topic by studying the heterogeneity in the willingness of consumers to share different types of personal data to different types of firms. We aim to quantify this heterogeneity. In particular, we have the following research questions:

1) *Are consumers willing to give consent to firms to use their data?*
2) *How does consumers' willingness to give firms access to their data depend on the following factors?*
    a. *the type of data;*
    b. *the type of firm;*
    c. *whether data are used anonymously or not;*
    d. *on financial incentives that firms provide?*

[1] European Commission (2020a). Data governance and data policies at the European Commission, accessed on 12 September 2021.
[2] European Commission (2020b). Digital Finance Strategy for the EU, accessed on 12 September 2021.
[3] See, e.g. https://www.oaic.gov.au/consumer-data-right/what-is-the-consumer-data-right/

3) *How does the dependence on these factors vary with the characteristics of consumers?*

Between 24 August 2020 and 6 September 2020, we conducted a survey among a representative panel of Dutch consumers to find the answers to these research questions. The survey included a discrete choice experiment to elicit how consumers' data sharing decisions depend on various attributes. A discrete choice experiment is a survey method where respondents are presented with hypothetical situations ('vignettes') that differ in several attributes. In our case the attributes of the hypothetical situation are the type of data, firm, anonymization and level of financial incentives. Consumers then have to choose between different situations. By having sufficient variation in choices within and between respondents, these choices allow for measurement of consumer data sharing preferences. Because respondents have to trade off different features of the vignettes simultaneously in realistic scenarios, vignettes allow for a more valid measurement of consumers' preferences compared to direct questioning.

Our study contributes in several ways to the existing literature on data sharing and privacy. First, we contribute to the literature on consumers' willingness to share personal data to firms. We examine the relative willingness to share different types of data with different types of firms. We find that consumers are more hesitant to share their data with webshops, BigTechs and insurers than they are to share their data with banks. We show that people are less likely to share health data and financial data on payments, wealth and pensions than they are to share other types of consumer data. Closest to our analysis in this respect is a paper by Prince and Wallsten (2020) who measure people's valuation of online privacy across six countries, a wide range of datatypes and various online platforms, using surveys with carefully designed choice sets of hypothetical vignettes. They focus on ten types of data people can share related to their mobile phone, payment account, and Facebook account. They find that across countries people attach the highest value on keeping information on their financial records and biometric data private. Prince and Wallsten also find substantial cross-country variation in how much value people attach to different types of data. In contrast, our paper takes a somewhat broader approach to the data types by focusing on classes of data. In addition, we include anonymity as a potential characteristic of how the data are shared and examines whether the way data are treated influences consumers' willingness to share data. Another related paper is Bijlsma et al. (2020). These authors research attitudes towards sharing payments data and find that that the propensity to give consent for payments data usage is highest if the data user is the own bank. Van der Cruijsen (2020) examines consumers' attitudes towards payments data usage by presenting them with different situations and asking them for each situation to what extent the use of payments data is acceptable. She finds that attitudes depend on the purpose of the data use. For example, most people support payments data usage to enhance safety. In contrast, support for commercial usage of payments data is very low, especially when the user is a firm other than the consumer's own bank.

By including anonymity as a potential characteristic of how data are shared, we also contribute to the literature on the role of anonymity in data sharing. Our results show that consumers are especially cautious in sharing their data when not used anonymously. Our work complements that of Benndorf and Normann (2018) and Regner and Riener (2017). Benndorf and Normann (2018) study the willingness to sell personal data in a laboratory setting. They find that subjects are almost always willing to sell anonymous data, in contrast to non-anonymous data, where one in six participants are not willing to sell personal information at all. Regner and Riener (2017) investigate the effect of reduced anonymity on consumers' purchase decisions (whether to buy, and if so how much to pay) at an online music store with "pay what you want" pricing and in an online experiment. They find that revealing customer information drastically reduced the number of purchasing customers. Hann et al. (2007) use a discrete choice experiment to quantify subjects' valuation of online privacy protection against improper access, error, and secondary use of personal information. They find that among US subjects, website privacy protection is worth \$30.49-\$44.62. With respect to payment instrument preferences, Van der Cruijsen and Van der Horst (2019) report, based upon survey results, that consumers find privacy an important payment instrument attribute. Acquisti at al. (2013) discuss the difference between the willingness to pay for a more privacy-protective offer and the willingness to accept a less privacy-protective offer. Their results highlight the sensitivity of privacy valuations to contextual factors. Also relevant is Bansal et al. (2016), who show that the extent to which an individual is prepared to disclose financial information to a finance website is positively related to the degree of trust in that website.

Third, we contribute to the literature on financial incentives in data sharing. We study the effect of financial incentives on the willingness to share data, using different levels of compensation, including no compensation at all. Our results show that financial rewards can trigger data sharing by part of the consumers. The effect levels off with the size of the reward and differs between consumer segments. Males, young people, highly educated people or people with a high income react stronger on the magnitude of the reward than others. In general, studies on the relationship between financial incentives and privacy have shown that it is hard to put a price on privacy (Acquisti et al. 2015). People tend to say they value privacy a lot, but are not very willing to pay for privacy (Acquisti et al. 2013). Regarding consumer behaviour in sharing information in a payments context, a particularly interesting study is the paper by Athey et al. (2017), who use data from a digital currency field experiment. They find that small changes in incentives, costs and information can have a significant influence on data sharing. Bijlsma et al. (2020) show that a financial incentive can tempt more people to use payments data related services, also when the service is offered by a firm other than one's own main bank. Again relevant is the work by Prince and Wallsten (2020) who find that privacy is relatively highly valued by women and people aged 45 and over. However, they do not find differences across income in privacy preferences as we consistently do.

Finally, our paper adds to the literature on heterogeneity in willingness to share data between people who differ in personal characteristics. In this respect, we are among the first that research different data types and pay special attention to trust and digital literacy, next to the standard demographic characteristics such as age, gender, income and education. We find that attitudes towards data usage depend on personal characteristics, consumers' digital skills and their trust in the firms using the data. Goldfarb and Tucker (2012) show that women and older individuals are more concerned with privacy issues than others. Using consumer survey data from the US, Armantier et al. (2021) find notable differences between demographic groups. Overall, US consumers have more trust in traditional financial institutions than government agencies or FinTechs with respect to safeguarding their personal data, and have the least trust in BigTechs. This pattern holds across demographic groups. However, there are differences in the level of trust. For example, people from racial minorities have less trust in financial institutions than non-Hispanic white people, while people aged 60 and over have lower trust in FinTechs and BigTechs than younger people. Bijlsma et al. (2020) find that the intended usage of new payments data-based services depends on trust in the providers of these services.

The remainder of the paper is organised as follows: Section 2 describes the set-up of our discrete choice experiment and our data. Section 3 provides descriptive results. Section 4 introduces the estimated model and the variables used in the data analysis. Section 5 presents and discusses the estimation results and Section 6 offers our conclusions.

2. The survey

We designed a unique survey to measure consumers' opinions regarding the privacy sensitivity of different types of data and their attitudes towards sharing these data with different types of firms and under different conditions (anonymity and financial compensation).

2.1 Data collection

We conducted the survey among 3,295 members of the CentERpanel between 24 August and 6 September 2020. It was fully completed by 2,483 of them (75%), and partially by 122 panel members (4%). Our analyses are based on the answers of 2,488 respondents. The CentERpanel is an online panel, managed by research institute CentERdata. It provides an accurate representation of the Dutch-speaking population in the Netherlands, aged 16 years and older.[4] In addition to the information collected in our survey, we use data on panel members' demographic

[4] For more information on the methodology, see Teppa and Vis (2012).

characteristics like age, gender and education. These characteristics are collected by CentERdata and are part of the annual DNB Household Survey (DHS).

2.2 Survey design

The survey starts with a question on respondants' actual sharing of payments data with different financial service providers during the past twelve months and the likelihood that they will share these data with them in the next twelve months. Here we distinguish between nine different service providers, i.e. the respondant's own bank where they hold their main payment account, other banks of which they are a customer, large technology firms like Apple, Facebook and Google, a webshop, a non-bank lender, a non-bank mortgage provider, a non-bank financial advisor, an insurance firm and other firms. Respondants could also indicate that (1) they had not given any firm permission to use their payments data, although they had received requests, or that (2) they had not given any firm permission to use their payments data, but were also not asked to do so. This part of the survey also contains a question on how much trust respondents have in the different (financial) service providers. We use this information to research whether data sharing decisions depend on trust in service providers. Next, the survey measures the privacy sensitivity of ten types of personal data (see Table 1), that can be valuable for different types of firms. Thereafter, the main body of our survey is presented to the respondents. It includes the vignettes that we use to measure consumers' attitudes towards sharing different types of their personal data with different types of firms under varying privacy and financial conditions. Here we mimic choice situations where we present respondents with different sets of choices that consumers in the Netherlands may already face, like sharing their payments data (PSD2), or which they may face in the near future when open data becomes a reality in Europe.

Table 1. Personal data

Category	Example/description
1. Payments data	ATM withdrawals, purchases, electronic payments
2. Wealth and debts	Income, pension, bank balance and debts
3. Personal characteristics	Gender, age, nationality, marital status, household composition, educational level, ethnicity, religion and sexual orientation
4. Contact details	Name, address, phone number, email
5 Health data	Visits to general practitioner (GP), medicine usage
6. Personal identification data	Citizen Service Number, passport number, ID-card number, driving license number and fingerprint
7. Geolocation data based on smartphone usage	Where you have been and when
8. Online search behaviour	Websites visited, videos watched, downloads
9. Social contacts	WhatsApp contacts, contacts other social media
10. Personal preferences	Media usage, political preferences, memberships of associations and sport clubs, hobbies

We exogenously vary the four attributes of interest in the vignettes and across the vignettes: type of personal data that is shared, type of service provider, compensation given by the service provider and anonymity. See Table 2 for an overview of the four attributes and their levels. To keep respondents motivated to make conscious choices, we limited the types of data to six and the type of service providers to four. The data types are: payments data, health data, location data, wealth data, personal data and data on preferences. We are interested to see how payments and wealth data are treated by consumers relative to other types of personal data. The four types of data receiving firms are: banks, insurance firms, large technology firms (BigTechs) and webshops. These firms are in the forefront of data sharing due to (upcoming) financial legislation, like PSD2, open finance and open data. BigTechs and webshops are also of interest as they already interact digitally with individuals, and data sharing may allow them to further enrich their databases with new information about existing and future customers. The attribute 'Financial compensation' concerns monthly payments from the service provider to data sharing individuals. This attribute can take on five values: No compensation, EUR 2, EUR 5, EUR 10 and EUR 20. We include a wide range of financial compensations as prior research shows that the amount people want to receive for sharing their data varies. The attribute 'Anonymity' captures the way personal data are processed and used by the service provider. We distinguish between anonymous and non-anonymous processing. In case of anonymous processing personal data cannot be traced back to the corresponding individual. In contrast, in case of non-anonymous data processing the service provider can link data to the corresponding individual, and use the personal data for instance for making customer specific offers.

Table 2. Attributes and levels used in the vignettes

Attributes	**Levels**
Type of data	1) Payments data, like ATM withdrawals, purchases and payments. 2) Health data, like General Practioner visits and medicine usage. 3) Location data from your smartphone, like where you have been and when. 4) Data on your wealth and pension. 5) Data on your personal characteristics, like your household composition, age and educational level. 6) Data on your personal preferences, like your hobbies, memberships and clothing style.
Data receiving firm	1) A bank 2) An insurer 3) A large technology firm 4) A webshop
Financial compensation	1) You will not receive a compensation 2) You will receive a monthly compensation of 2 euros for this 3) You will receive a monthly compensation of 5 euros for this 4) You will receive a monthly compensation of 10 euros for this 5) You will receive a monthly compensation of 20 euros for this
Anonymity	1) Your data will **not** be anonymized. 2) Your data will be anonymized.

In total there are 240 different hypothetical data sharing situations (6*4*5*2) and 28,680 different two-choice vignettes (240*239/2). We selected a subset of all possible choice sets, using a statistical STATA software routine called *dcreate* by Hole (2016) that constructs a fractional factorial D-optimal design. A D-optimal design varies the levels of each attribute for each choice and for each respondent in such a way that with a limited number of choice sets the influence of the different attributes on individuals' choices is estimated as precisely as possible (for more information, see e.g. Carlsson and Martinsson 2003; Zwerina et al. 1996). We chose a design which generates 2,400 vignettes with two alternatives. Our relative D-efficiency is 76.9%. We grouped the resulting vignettes into 240 sets of ten and randomly distributed these sets across our respondents. So, every panel member was randomly assigned to one of the 240 subsets, each consisting of ten pairs of choice sets, i.e. the vignettes.

The set of repeated choices was introduced as follows: *"Suppose two different types of firms, such as a bank and a large technology firm (e.g. Apple, Facebook or Google) ask you to share data with them. This way they can serve you better, for example by helping you faster and offering you better products. The type of data that firms ask you to share with them may be different. You can think for example of data about your health, your finances or your geographical location. Firms can give you compensation for sharing your data, but they do not have to do so. Some firms anonymize your data, so that it cannot be traced back to you, whereas other firms don't. You will now be presented with 10 situations. Please indicate which of the two different types of data sharing you prefer. You may not prefer either option—nevertheless, we still ask you to make a choice. An example is shown below for illustration."* Figure 1 is an example of how the first vignette was presented to the panel members.

Figure 1. Example of a vignette

Option 1	**Option 2**
An insurance firm wants to receive payments data, such as withdrawals, purchases and payments. You will not receive a monthly compensation. Your data will be anonymized.	A bank wants to receive data on your personal characteristics, like your gender, household composition, age and educational level. You will receive a monthly compensation of two euros. Your data will **not** be anonymized.

Which option do you choose?

Option 1 ○ **Option 2** ○

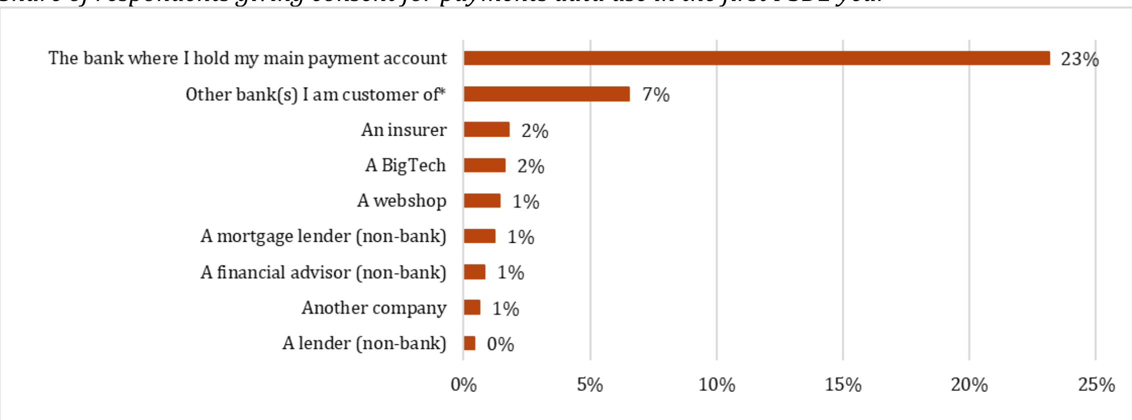
The survey ends with questions on people's self-assessed level of digital skills, their online shopping behaviour and social media usage. We use the answers to these questions to research whether data sharing decisions depend on digital skills, online shopping behaviour and social media usage.

3. Survey outcomes: descriptive statistics

3.1 Data sharing and privacy

First, we look at respondents' actual payments data sharing behaviour with different firms between August 2019 and August 2020. We asked the following question: *"In the past twelve months, which of the following firms did you give permission to use the payments data of your main payment account to offer services? For example, services like an app that gives an overview of income and expenses, providing a loan, or to help you with budget management."* A quarter of the respondents indicated that they authorised the use of their payments data to use new payment related services in the first year in which PSD2 was in force in the Netherlands. They predominantly authorized the banks with which they have their main current account[5] to access their payments data, followed by other banks where they hold an account (Figure 2). A small part of the respondents stated they also granted access to other (licensed) firms. For example, 2% indicated they allowed insurance firms and BigTechs– such as Apple, Facebook and Google – access to their payments data and 1% allowed webshops to use their data. Of the 75% of the respondents who had not given permission to any firm to use their payments data, 14% said they were asked to do so, but decided not to do it, and 86% said they were not asked permission to use their payments data.

Figure 2. Consumers predominantly authorize own bank to use payments data
Share of respondents giving consent for payments data use in the first PSD2 year



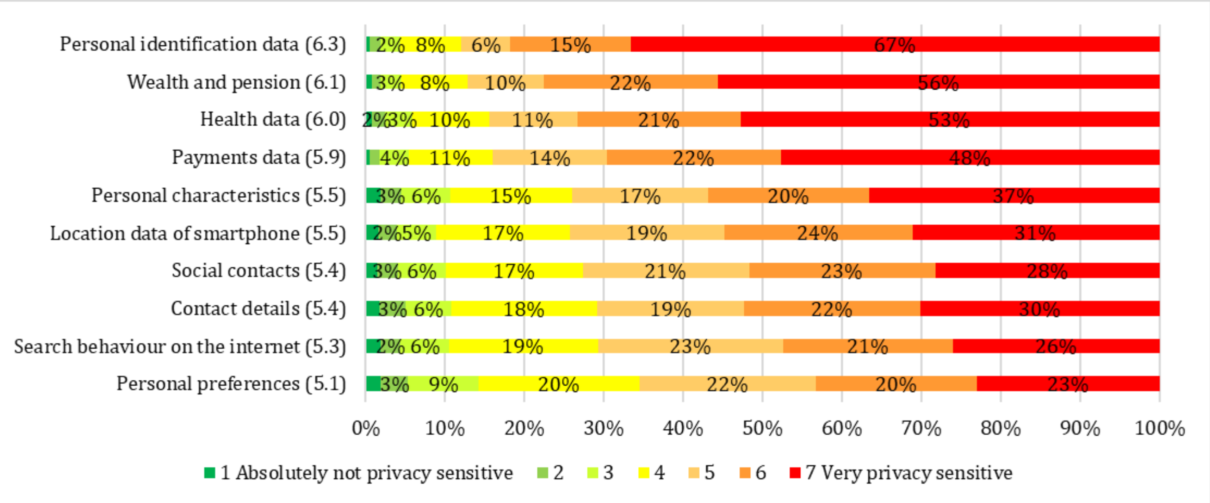
Note: Respondents indicated for each service provider whether they gave consent. 2,488 respondents. *Only answered by 1,160 respondents with accounts at multiple banks.

[5] Note that BigTechs were not licensed (yet) for PSD2 services at the time the survey was held. Dutch consumers could not give these firms access to their payment account data yet, as intended by PSD2. Maybe respondents who used a mobile payment app that banks offer in co-operation with technology firms stated they shared payments data with a technology firms (Samsung pay, Google pay, etc).

Second, respondents were asked about the likelihood that they would give permission to licensed firms to use their payments data in exchange of services in the next twelve months. The question reads as follows: *"What is the likelihood that you would give the following parties within the next twelve months permission to use the payments data of your main payment account to offer services? Fill in a number between 0 and 100 (0 = I will definitely not give permission and 100 = I will certainly give permission)."* Again, the average likelihood that a firm would get permission from the respondents is highest for banks where respondents have their main payment account (26%), followed by other banks they are already customers of (11%). The likelihood that they would give a mortgage lender or financial advisor access to their payments data is 4%, and that they would give it to an insurance firm is 3%. The likelihood is lowest for webshops, BigTechs, banks they are not customers of and lenders (in all cases: 2%). 53% of respondents indicated a probability of 0% for all providers. These respondents definitely do not want to give permission to any party.

Third, we consider the privacy sensitivity of certain data types. Respondents were asked to assess the privacy sensitivity of ten different data types (see Table 1). The question reads as follows: "*How privacy sensitive do you find the following types of data? Please give a number from 1 to 7, where 1 stands for "not at all privacy sensitive" and 7 for "very privacy sensitive".*

Figure 3. Financial data are perceived as very privacy sensitive
Response shares

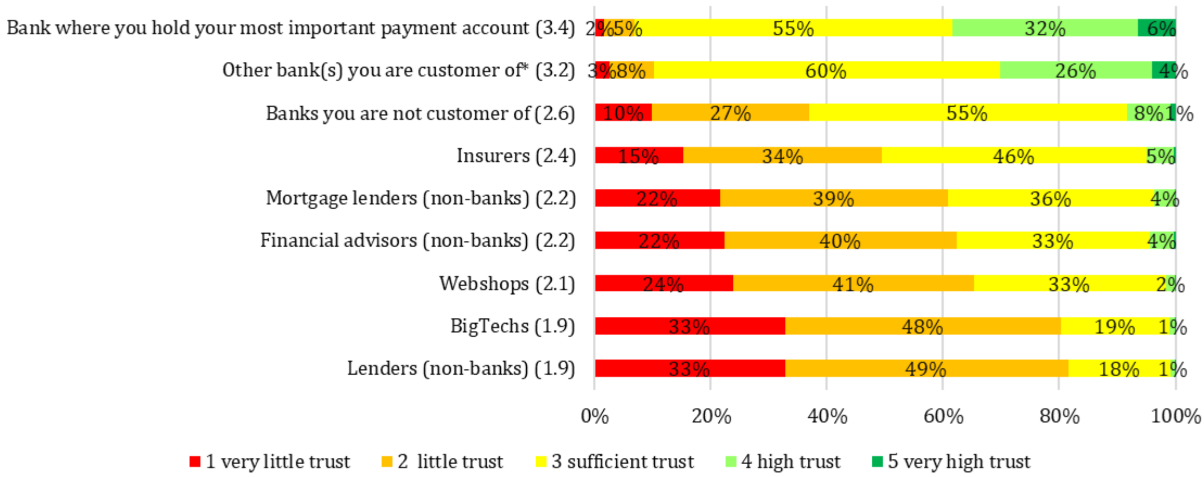


Note: 2,488 respondents. The average privacy sensitivity is in brackets behind the type of data.

The average privacy assessments range between 5.1 and 6.3, indicating that the respondents perceive all listed types of personal data as privacy sensitive (see Figure 3). Personal identification data are considered as the most privacy sensitive type of information. Financial data, such as data on wealth and pension and data on payment transactions and cash withdrawals are also perceived as very privacy sensitive, as well as health data. Consumers find these types of data more privacy sensitive than information on their internet search behaviour, their social contacts, the location data of their smartphones, their contact details and data on their personal preferences.[6]

Fourth, we consider how much trust respondents have in different service providers, with whom they may share their personal data and find that they have most trust in their own bank. Respondents were asked the following question: *"How much trust do you have in [name service provider]?"*, using a 1 (very little trust) to 5-point scale (very high trust). The average trust assessments range between 1.9 and 3.4 (see Figure 4). Only banks where the respondent is customer of score on average above 3, indicating that respondents trust them most. Other banks of which they are not customers and insurance firms get an average trust rating of 2.6 and 2.4 respectively, indicating that people have less trust in them. Respondents trust BigTechs and non-bank lenders least, they get on average a score below 2.[7]

Figure 4. People have most trust in their own main bank
Response shares



Note: 2,488 observations. The average score is provided between brackets. *Only answered by 1,160 respondents with accounts at multiple banks.

3.2 Vignettes

Our primary focus in this paper is the analysis of the decisions made by the respondents in the discrete choice experiment. Table 3 summarizes the distribution of the choices made by the respondents over the different levels of the data receiving firms, the type of data to be shared, the monthly compensation and the way the data are processed (anonymously or not). We show the results for the whole sample (column 1) and by gender (columns 2 and 3), age group (columns 4-6), educational level (columns 7 and 8) and income group (columns 9-11). Below, we make some

[6] Using two-sided t-tests we tested whether respondents perceive the privacy sensitivity of the 10 data classes as equal or not. They consider the privacy sensitivity of most of the 10 data types as different (p<0.01). They only perceive information about their social contacts and their contact details as equally sensitive (p=0.67) as well as information about their personal characteristics and the location data of their smartphone (p=0.11).

[7] The relative trust ranking of banks, insurance firms and BigTechs of the Dutch is in line with the relative trust ranking of US citizens, see Van der Cruijsen et al. (2021) and Armantier et al. (2021).

initial observations based on the descriptive statistics in Table 3. Of course, these observations need to be analysed by estimating a choice model.

Table 3. Breakdown of the characteristics of the choices made by gender, age, income and education

	(1) Whole sample	(2) Female	(3) Male	(4) Age ≤34	(5) Age 35-54	(6) Age ≥55	(7) Education low or medium	(8) Education high	(9) Income low	(10) Income middle	(11) Income high
Payments data	14%	13%	14%	14%	13%	14%	14%	14%	13%	14%	14%
Health data	13%	14%	12%	14%	13%	13%	13%	12%	14%	13%	12%
Location data smartphone	17%	17%	17%	14%	17%	17%	17%	17%	16%	17%	17%
Wealth and pensions	15%	15%	15%	15%	15%	15%	15%	15%	15%	14%	16%
Personal characteristics	20%	20%	20%	20%	21%	20%	20%	21%	20%	20%	21%
Personal preferences	21%	21%	21%	23%	21%	20%	21%	21%	21%	21%	20%
Bank	29%	29%	29%	28%	29%	29%	29%	29%	29%	29%	29%
Insurer	26%	26%	26%	26%	25%	26%	26%	26%	26%	26%	26%
BigTech	23%	24%	23%	24%	23%	23%	23%	23%	23%	23%	23%
Webshop	22%	22%	21%	22%	22%	21%	22%	22%	22%	22%	22%
0 euro	18%	19%	18%	17%	17%	19%	19%	18%	19%	18%	19%
2 euros	20%	20%	19%	19%	19%	20%	20%	19%	20%	20%	19%
5 euros	20%	20%	20%	19%	20%	20%	20%	19%	20%	20%	20%
10 euros	21%	21%	20%	21%	21%	20%	21%	21%	20%	20%	21%
20 euros	22%	21%	22%	23%	23%	21%	21%	22%	21%	22%	22%
Not anonymous	28%	28%	27%	28%	25%	29%	30%	23%	30%	30%	24%
Anonymous	72%	72%	73%	72%	75%	71%	70%	77%	70%	70%	76%
Number of vignettes	24,767	11,968	12,799	3,093	7,581	14,093	15,377	9,370	5,376	8,703	9,348

Note: The respondents made 24,767 binary choices. The table reports the share of the four firm types, six data types, five levels of monthly reward and two types of data processing in all resulting choices made by all respondents in the sample and by gender, age category, income category and educational level.

This first breakdown suggests that the respondents are most keen on sharing their data with banks (29% of choices), and least with webshops (22%). Insurers rank second (26%) and BigTechs third (23%). This ordering holds for all demographic groups. The breakdown also suggests that respondents find an anonymous way of data processing much more attractive than non-anonymous data usage. Respondents select an anonymous way of processing of their data in 72% of their choices. People aged between 35 and 54, with a high household income or with at least a bachelor degree choose relatively more often for anonymous processing of their data than others. Last, respondents are sensitive to rewards. The share of being selected in the offered choices rises from 18% if no financial compensation is offered by the data receiving firm to 22% if monthly financial compensation of 20 euros is offered. The 18% share in case of no financial compensation suggests that for many people, factors other than money may be more important when deciding to share data or not. In addition, we see that sensitivity to incentives differs by gender, age, educational level and income. Overall, males, people aged 54 and younger, people with a medium to high income or with at least a bachelor degree react more strongly to financial rewards than others.

Table 4 provides an overview of the average value of the rewards for the choices made by the respondents in the discrete choice experiment. The average value of the reward over all choices made is EUR 7.76/month (column 11), The average reward varies between EUR 6.09/month for insurance firms that receive information on personal characteristics and process these data in an anonymous way (column 4) and EUR 9.75/month for insurance firms that receive information on personal preferences and link these data to the individuals (column 3).

Table 4. Average monthly reward for the choices made (in euros)

	Bank		Insurer		BigTech		Webshop		All firm types		All
	(1) Not anony-mous	(2) Anony-mous	(3) Not anony-mous	(4) Anony-mous	(5) Not anony-mous	(6) Anony-mous	(7) Not anony-mous	(8) Anony-mous	(9) Not anony-mous	(10) Anony-mous	(11)
Payments data	6.60	7.55	8.22	8.12	7.78	8.38	9.32	7.73	7.71	7.94	7.89
	302	*775*	*160*	*796*	*154*	*584*	*150*	*534*	*766*	*2,689*	*3,455*
Health data	8.39	8.37	7.65	9.30	7.85	6.56	7.16	7.83	7.84	8.00	7.97
	210	*659*	*238*	*605*	*149*	*640*	*107*	*606*	*704*	*2,510*	*3,214*
Location data smartphone	7.49	7.49	7.21	6.66	8.48	8.43	7.61	8.15	7.64	7.69	7.67
	361	*829*	*299*	*726*	*232*	*775*	*253*	*718*	*1,145*	*3,048*	*4,193*
Wealth and pensions	8.59	8.39	7.39	6.81	8.18	7.72	7.57	8.45	8.06	7.79	7.85
	367	*871*	*226*	*833*	*164*	*620*	*129*	*504*	*886*	*2,828*	*3,714*
Personal characteristics	6.98	6.96	8.47	6.09	8.93	7.49	7.09	7.53	7.88	7.07	7.33
	453	*1,016*	*469*	*658*	*360*	*872*	*313*	*867*	*1,595*	*3,413*	*5,008*
Personal preferences	9.55	8.65	9.75	7.54	8.64	6.60	7.69	7.18	8.97	7.49	7.98
	531	*805*	*394*	*1,004*	*417*	*807*	*378*	*847*	*1,720*	*3,463*	*5,183*
Total	8.02	7.85	8.27	7.39	8.47	7.51	7.66	7.18	8.11	7.63	7.76
Number of vignettes	2,224	4,955	1,786	4,622	1,476	4,298	1,330	4,076	6,816	17,951	24,767

Note: The table presents average monthly rewards, expressed in euros for each combination of data type, firm type and way of data processing. The averages correspond to the group averages for choices made by the respondents. The numbers in italics present the number of choices per combination.

4. Methodology

We estimate conditional logit models. These models are appropriate to model the choice among alternatives as a function of characteristics of these alternatives. Equation 1 is a linear random utility model. $\boldsymbol{x}_{ijk}$ is a vector of attributes – the type of data, the type of firm, the financial compensation and the anonymity of data usage – for alternative j in the vignette k that individual i faces.

$$u_{ijk} = \boldsymbol{x}'_{ijk}\boldsymbol{\beta} + \varepsilon_{ijk} \quad (1)$$

With the assumption that ε_{ijk} is independently and identically distributed with type I extreme value distributions the probability that individual i chooses data usage alternative j among two alternatives in vignette k is given in equation (2).

$$Prob(Y_{ik} = j) = \frac{\exp(\boldsymbol{x}'_{ijk}\boldsymbol{\beta})}{\sum_{n=1}^{2}\exp(\boldsymbol{x}'_{ink}\boldsymbol{\beta})} \quad (2)$$

For all vignettes we know which of the two data usage alternatives respondents chose. Therefore, we can generate the likelihood function based on the probabilities. The likelihood function is optimized with respect to $\boldsymbol{\beta}$ and the estimated utitlity parameters for each attribute are obtained, while the errors are clustered on individuals.

The set of attributes consists of dummy variables capturing the type of data, type of firm, financial compensation and anonymity of data usage. The variables capturing the data type are: *health data*, *location data smartphone*, *wealth and pensions*, *personal characteristics*, and *personal preferences*. The reference category is *payments data*. For example, *health data* is 1 for options in which the data type is health data. For example, *insurer* is 1 in case the type of firm in an option is an insurer and 0 in case another firm was included in the option. In a similar fashion *BigTech* and *webshop* are constructed. The reference category is an option in which a *bank* uses the data. *2 euros*, *5 euros*, *10 euros*, and *20 euros* capture the financial compensation in the option. The reference category is *no compensation*, so an option without a financial reward. To calculate the willingness-to-accept (WTA) for attributes, we estimate conditional logit models with *reward* included as a continuous variable instead of the reward dummies and rely on $\boldsymbol{\beta}$. The variable *anonymous* is 1 for the option with anonymous data sharing and 0 for options with a non-anonymous way of data sharing.

We expect that the likelihood of data sharing depends positively on the financial reward and the data being treated anonymously. As respondents indicated that they find data on wealth and pensions, health and payments to be the most sensitive, we anticipate that they are least likely to share these data types. Based on our findings on trust in the different service providers, we expect that consumers are more likely to share their data with banks than with insurers, BigTechs and webshops.

5. Regression results

5.1 Data sharing depends on the data type, data user, compensation and anonymity

The results of conditional logit regressions show that data sharing choices depend on the type of firm, the type of data, the financial compensation and whether the data are used anonymously. Column 1 of Table 5 shows the results for the whole sample. Table 5 also shows the results for different subgroups based on gender (column 2 and 3), age (column 4 and 5), the level of education (6 and 7) and income (column 8, 9 and 10).

Table 5. Regression results: demographic groups

	(1) All	(2) Men	(3) Women	(4) Age <45	(5) Age ≥45	(6) Low education	(7) High education	(8) Low income	(9) Middle income	(10) High income
Data type (reference category: payments data)										
Health data	-0.03***	-0.05***	-0.01	-0.01	-0.04***	-0.01	-0.06***	0.01	-0.02	-0.07***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)
Location data smartphone	0.09***	0.08***	0.10***	0.06***	0.09***	0.08***	0.10***	0.08***	0.09***	0.08***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)
Wealth and pensions	0.03***	0.02**	0.03***	0.05***	0.02**	0.03***	0.03**	0.04***	0.02	0.03**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)
Personal characteristics	0.17***	0.15***	0.19***	0.18***	0.17***	0.17***	0.18***	0.18***	0.17***	0.16***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Personal preferences	0.18***	0.16***	0.21***	0.20***	0.18***	0.19***	0.18***	0.20***	0.18***	0.17***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)
Firm (reference category: bank)										
Insurer	-0.05***	-0.05***	-0.05***	-0.04***	-0.05***	-0.05***	-0.05***	-0.05***	-0.05***	-0.06***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
BigTech	-0.09***	-0.11***	-0.08***	-0.08***	-0.10***	-0.10***	-0.09***	-0.08***	-0.10***	-0.11***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Webshop	-0.12***	-0.13***	-0.10***	-0.10***	-0.13***	-0.12***	-0.11***	-0.10***	-0.12***	-0.13***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Financial compensation (reference category: no compensation)										
2 euros	0.02***	0.02**	0.03***	0.04***	0.02**	0.02**	0.03***	0.01	0.04***	0.02*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
5 euros	0.03***	0.04***	0.02**	0.04***	0.03***	0.03***	0.03***	0.02*	0.02**	0.05***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
10 euros	0.05***	0.06***	0.04***	0.08***	0.04***	0.05***	0.06***	0.03**	0.05***	0.07***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
20 euros	0.07***	0.09***	0.04***	0.11***	0.05***	0.05***	0.09***	0.05***	0.07***	0.08***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Data processing (reference category: non-anonymous)										
Anonymous	0.22***	0.23***	0.21***	0.22***	0.22***	0.19***	0.26***	0.19***	0.20***	0.26***
	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Number of vignettes	24,767	12,799	11,968	6,485	18,282	15,377	9,370	5,376	8,703	9,348
Pseudo R-squared	0.22	0.23	0.22	0.25	0.22	0.18	0.31	0.19	0.19	0.29
Log pseudolikelihood	-13,343.7	-6,822.7	-6,482.5	-3,369.6	-9,941.9	-8,713.5	-4,498.4	-3,034.8	-4,907.4	-4,589.6
Wald $\chi 2$	2441.8***	1,396.8***	1,070.4***	855.0***	1,689.3***	1,322.2***	1,364.7***	436.6***	799.0***	1,242.8***

Note: The table reports average marginal effects for conditional logit regressions. Standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Regarding the six different data types, people are least likely to share their health data, followed by payments data. The likelihood of sharing health data is 3 percentage points (p.p). lower than the likelihood of sharing payments data. For all other data types, it holds that people are more willing to share these than payments data. Compared to payments data, people are 3 p.p. more likely to share data about wealth and pensions and 9 p.p. more likely to share location data from the smartphone. Consumers are most likely to opt for data usage on personal preferences and data on personal characteristics. The likelihood of sharing these two data types is 18 p.p. and 17 p.p. higher than the likelihood of sharing payments data.

There are a few differences in the ranking of the likelihood of sharing different data types between different groups of people. Men, people aged 45 and over, highly educated people and high-income individuals are least likely to share their health data. In contrast, women, people under 45, less educated people and people with a low or medium income are as likely to share their health data as they are to share their payments data. Regressions for more detailed age classes show that people younger than 25 are as likely to share smartphone location data and data on wealth and pensions as they are to share data on their health and payments data (see Table A.1 in Appendix A). For people with a medium income, the likelihood of sharing data on wealth and pensions does not significantly differ from the likelihood of sharing payments data and the likelihood of sharing health data.

Dutch consumers are more likely to give their consent for data usage by banks than for usage by other types of firms. Compared to banks, they are 5 percentage points (p.p.) less likely to give consent to insurers, 9 p.p. less likely to agree with data usage by BigTechs and 11 p.p. less likely to give their approval to webshops. The gap in the likelihood of agreeing with data usage is higher for men than for women in case of BigTechs and webshops. In case of insurers there is no gender difference. The difference in likelihood of sharing data with banks compared to other firms is highest for people aged 45 and above and for high-income people.

There is a positive relationship between the level of financial compensation offered and the likelihood of agreeing with the data usage. When the financial compensation is 2 euros per month, people are 2 p.p. more likely to agree than if there were no financial compensation. In case of 5, 10 and 20 euros these effects are respectively 3, 5 and 7 p.p.. The effect of financial compensation on the likelihood of data sharing is therefore non-linear; the marginal impact of increasing compensation reduces with the level of compensation. Men, young people, people with a high level of education and people with a high level of income are more sensitive to financial compensation than women, old people, less educated people and low-income people. For people aged 65 or older we find that the likelihood of data sharing is higher when 2 euros is being offered but unaltered when more compensation is being given (Table A.1 in Appendix A).

When data usage is anonymous the likelihood that people consent to data usage strongly increases. The likelihood of agreeing to the data usage is 22 p.p. higher when the data are used anonymously than when they arep not used anonymously, i.e. that they can be linked to individuals. The effect of anonymity on the likelihood of sharing data is relatively high for men, high-income and highly educated people.

5.2 Data sharing depends on digital skills, webshop usage and social media usage

We also examined whether differences in digital skills, as reflected by people's digital literacy and the extent in which they are active online, as reflected by people's webshop usage and social media usage, influence data sharing. It may be possible that respondents with high digital skills, who do a lot of online shopping or who are active on social media platforms differ in the kind of data they prefer (not) to share, and with whom they would like to share data compared to other people. We ran regressions for different subgroups of people based on their digital skills.

First, we distinguish between people with low digital literacy and people with high digital literacy. Respondents who say they agree or fully agree with the statement "I can work well with a computer, tablet and smartphone" are in the high digital literacy subgroup. Respondents who disagree or take a neutral stance are in the low digital literacy subgroup. The regression results of the low and high digital literacy groups are in respectively Table 6 column 2 and 3. Second, we make three groups based on monthly webshop usage prior to the survey. The results of respondents who (1) did not use webshops, (2) used webshops 1-4 times, and (3) used webshops 5 times or more are in Table 6 column 4, 5 and 6. Last, we separate respondents based on their social media usage. Table 6, column 7 shows the results for respondents who never use social media such as Instagram, WhatsApp, Facebook, Twitter or YouTube. The results for respondents who use social media at most once a day are in column 8, whereas the findings on more frequent users are in column 9.

We find that attitudes towards data sharing depend on people's digital skills, their social media usage and online shopping behaviour. People who do not use social media are more likely to share data with banks and unlikely to share data with other firms than people who are active on social media. People who do not visit webshops are less likely to want to share data with webshops than people who use webshops. People using social media or webshops may be more used to sharing data with other people or firms than people who are less active online. Of course it may also be that the latter group just do not see themselves coming into the position of sharing their data with other parties than their own bank. The ranking of different data types based on the likelihood of sharing the data does not differ much between people with low and high digital literacy. The only difference is that people with high digital literacy are less likely to share their health data than their payments data, whereas the likelihood of sharing payments data and health

Table 6. Regression results: digital skills, webshop usage and social media usage

	(1) All	(2) Low digital literacy	(3) High digital literacy	(4) Webshop usage: no	(5) Webshop usage: low	(6) Webshop usage: high	(7) Social media usage: no	(8) Social media usage: low	(9) Social media usage: high
Data type (reference category: payments data)									
Health data	-0.03***	-0.01	-0.04***	0.01	-0.04***	-0.06**	-0.02	-0.02	-0.03***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.01)
Location data smartphone	0.09***	0.08***	0.09***	0.08***	0.09***	0.10***	0.06**	0.09***	0.09***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)	(0.03)	(0.01)	(0.01)
Wealth and pensions	0.03***	0.02*	0.03***	0.03*	0.02**	0.05**	0.00	0.02*	0.03***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)	(0.03)	(0.01)	(0.01)
Personal characteristics	0.17***	0.16***	0.18***	0.15***	0.17***	0.19***	0.16***	0.15***	0.18***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)
Personal preferences	0.18***	0.17***	0.19***	0.17***	0.18***	0.20***	0.17***	0.17***	0.19***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)
Firm (reference category: bank)									
Insurer	-0.05***	-0.05***	-0.05***	-0.07***	-0.04***	-0.06***	-0.08***	-0.04***	-0.05***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)
BigTech	-0.09***	-0.08***	-0.10***	-0.10***	-0.09***	-0.10***	-0.15***	-0.07***	-0.10***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)
Webshop	-0.12***	-0.12***	-0.12***	-0.15***	-0.10***	-0.12***	-0.17***	-0.11***	-0.11***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)
Financial compensation (reference category: no compensation)									
2 euros	0.02***	0.02*	0.02***	-0.01	0.04***	0.03*	0.03	0.01	0.03***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)
5 euros	0.03***	0.02**	0.04***	-0.00	0.05***	0.04**	0.01	0.02**	0.04***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)
10 euros	0.05***	0.03**	0.07***	0.03**	0.06***	0.08***	0.04	0.03**	0.07***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)
20 euros	0.07***	0.03***	0.08***	0.01	0.09***	0.08***	0.03	0.03***	0.09***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.03)	(0.01)	(0.01)
Data processing (reference category: non-anonymous)									
Anonymous	0.22***	0.20***	0.23***	0.20***	0.23***	0.23***	0.18***	0.21***	0.23***
	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)
Number of vignettes	24,767	8,570	16,150	6,890	14,970	2,860	2,240	7,800	14,680
Pseudo R-squared	0.22	0.17	0.25	0.16	0.25	0.26	0.16	0.20	0.25
Log pseudolikelihood	-13,343.7	-4,925.8	-8,343.0	-3,993.9	-7,772.6	-1,469.8	-1,308.1	-4,342.2	-7,610.2
Wald $\chi 2$	2,441.8***	618.7***	2,090.7***	495.5***	1,771.0***	401.1***	196.2***	642.1***	1,720.9***

Note: The table reports average marginal effects for conditional logit regressions. Standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

data is the same for people with a low level of digital literacy.

People who do not use webshops are as likely to share payments data as health data. The same holds for people who use social media rarely or never. For the other subgroups based on social media and webshop usage we find a lower likelihood of sharing health data than of sharing payments data. For people who do not use social media the likelihood of sharing data on wealth and pensions does not differ significantly from the likelihood of sharing data on health and payments. People with high digital literacy, frequent webshop users and social media users are less hesitant to share data on personal characteristics and preferences than other groups of people. Also, people with high digital literacy, people who use webshops and frequent social media users are more sensitive to financial compensation than people with lower digital literacy, people who do not use webshops and people who do not use social media or use it less frequently. The effect of anonymity on the likelihood of data sharing is relatively high for people with high digital skills, webshop users and social media users. This suggests these users may be more aware of the relevance of anonymity in protecting their privacy.

5.3 Data sharing depends on trust in the data using firm

Next, we examine whether data sharing decisions depend on trust in the firm that wants to use the data in exchange for delivering a service. Table 7 shows the results of regressions for different subgroups of respondents based on their trust in their own main bank, other banks they are not customer of, insurers, BigTechs, and webshops. We separate respondents for each type of trust in these firms into three groups: (1) people with little or very little trust, (2) people with sufficient trust and (3) people with high or very high trust.

We find that the likelihood of sharing data with a particular type of firm depends on people's trust in that type of firm. Compared to people with little trust in banks, people with high trust in banks are more likely to share their data with banks than with BigTechs or webshops. For example, someone with high trust in their own bank is 15 p.p. less likely to share data with webshops than with banks, whereas someone with low trust in the own main bank is only 7 p.p. less likely to share data with webshops than with banks. These effects are 11 p.p. and 6 p.p. in case of BigTechs. These are substantial and significant differences. We also find that people with high trust in insurers do not report a significant difference in the likelihood of sharing data with banks and insurers.

Table 7. Regression results: trust levels (1/2)

	(1) All	(2) Own main bank: low trust	(3) Own main bank: sufficient trust	(4) Own main bank: high trust	(5) Non-own bank: low trust	(6) Non-own bank: sufficient trust	(7) Non-own bank: high trust	(8) Insurers: low trust	(9) Insurers: sufficient trust	(10) Insurers: high trust
Data type (reference category: payments data)										
Health data	-0.03***	-0.05*	-0.03**	-0.03*	-0.01	-0.05***	-0.01	-0.04***	-0.02*	-0.02
	(0.01)	(0.03)	(0.01)	(0.01)	(0.01)	(0.01)	(0.03)	(0.01)	(0.01)	(0.04)
Location data smartphone	0.09***	0.02	0.10***	0.08***	0.08***	0.09***	0.12***	0.08***	0.09***	0.08**
	(0.01)	(0.03)	(0.01)	(0.01)	(0.01)	(0.01)	(0.03)	(0.01)	(0.01)	(0.04)
Wealth and pensions	0.03***	0.01	0.02**	0.04***	0.03*	0.02*	0.10***	0.02*	0.04***	0.06*
	(0.01)	(0.03)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.03)
Personal characteristics	0.17***	0.09***	0.17***	0.19***	0.16***	0.18***	0.21***	0.15***	0.19***	0.19***
	(0.01)	(0.03)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.03)
Personal preferences	0.18***	0.12***	0.18***	0.20***	0.17***	0.19***	0.21***	0.16***	0.21***	0.16***
	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.04)
Firm (reference category: bank)										
Insurer	-0.05***	-0.05**	-0.04***	-0.06***	-0.05***	-0.05***	-0.03*	-0.05***	-0.05***	-0.02
	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)
BigTech	-0.09***	-0.06***	-0.09***	-0.11***	-0.09***	-0.09***	-0.10***	-0.09***	-0.10***	-0.09***
	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.03)
Webshop	-0.12***	-0.07***	-0.11***	-0.15***	-0.10***	-0.12***	-0.15***	-0.11***	-0.12***	-0.15***
	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.03)
Financial compensation (reference category: no compensation)										
2 euros	0.02***	0.04*	0.02***	0.02*	0.03**	0.02**	0.02	0.01	0.03***	0.06**
	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.03)
5 euros	0.03***	0.06**	0.03***	0.04***	0.03***	0.03***	0.06***	0.04***	0.03***	0.07**
	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.03)
10 euros	0.05***	0.08***	0.04***	0.06***	0.04***	0.06***	0.07***	0.05***	0.05***	0.08***
	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.03)
20 euros	0.07***	0.09***	0.06***	0.08***	0.05***	0.07***	0.10***	0.06***	0.07***	0.09***
	(0.01)	(0.03)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.03)
Data processing (reference category: non-anonymous)										
Anonymous	0.22***	0.23***	0.22***	0.21***	0.20***	0.24***	0.20***	0.22***	0.22***	0.19***
	(0.00)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
Number of vignettes	24,767	1,711	13,530	9,526	9,131	13,573	2,063	12,224	11,333	1,210
Pseudo R-squared	0.22	0.21	0.23	0.22	0.18	0.26	0.24	0.21	0.24	0.21
Log pseudolikelihood	-13,343.7	-939.4	-7,250.5	-5,120.4	-5,196.2	-7,000.3	-1,090.6	-6,670.3	-5,979.2	-662.9
Wald $\chi 2$	2,441.8***	157.4***	1,275.3***	1,063.7***	714.1***	1,602.5***	298.8***	1,065.5***	1,301.4***	197.3***

Note: The table reports average marginal effects for conditional logit regressions. Standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 7. Regression results: trust levels (2/2)

	(1) All	(2) BigTechs: low trust	(3) BigTechs: sufficient trust	(4) BigTechs: high trust	(5) Webshops: low trust	(6) Webshops: sufficient trust	(7) Webshops: high trust
Data type (reference category: payments data)							
Health data	-0.03***	-0.03***	-0.01	-0.10	-0.03***	-0.02	-0.07
	(0.01)	(0.01)	(0.02)	(0.08)	(0.01)	(0.01)	(0.07)
Location data smartphone	0.09***	0.10***	0.06***	0.04	0.09***	0.09***	0.04
	(0.01)	(0.01)	(0.02)	(0.05)	(0.01)	(0.01)	(0.06)
Wealth and pensions	0.03***	0.03***	0.03	-0.01	0.02**	0.05***	-0.08
	(0.01)	(0.01)	(0.02)	(0.06)	(0.01)	(0.01)	(0.07)
Personal characteristics	0.17***	0.17***	0.17***	0.18***	0.17***	0.18***	0.16***
	(0.01)	(0.01)	(0.02)	(0.05)	(0.01)	(0.01)	(0.05)
Personal preferences	0.18***	0.18***	0.19***	0.15**	0.18***	0.19***	0.12*
	(0.01)	(0.01)	(0.02)	(0.06)	(0.01)	(0.01)	(0.06)
Firm (reference category: bank)							
Insurer	-0.05***	-0.05***	-0.05***	-0.07*	-0.05***	-0.04***	-0.05
	(0.01)	(0.01)	(0.01)	(0.05)	(0.01)	(0.01)	(0.04)
BigTech	-0.09***	-0.10***	-0.07***	-0.04	-0.10***	-0.08***	-0.15***
	(0.01)	(0.01)	(0.01)	(0.05)	(0.01)	(0.01)	(0.05)
Webshop	-0.12***	-0.12***	-0.10***	-0.13***	-0.13***	-0.09***	-0.17***
	(0.01)	(0.01)	(0.01)	(0.06)	(0.01)	(0.01)	(0.04)
Financial compensation (reference category: no compensation)							
2 euros	0.02***	0.02***	0.03**	0.18***	0.01	0.05***	0.09
	(0.01)	(0.01)	(0.01)	(0.06)	(0.01)	(0.01)	(0.05)
5 euros	0.03***	0.03***	0.04**	0.12*	0.02***	0.05***	0.05
	(0.01)	(0.01)	(0.01)	(0.07)	(0.01)	(0.01)	(0.05)
10 euros	0.05***	0.05***	0.07***	0.20**	0.04***	0.08***	0.10**
	(0.01)	(0.01)	(0.01)	(0.08)	(0.01)	(0.01)	(0.04)
20 euros	0.07***	0.06***	0.10***	0.29***	0.05***	0.09***	0.16***
	(0.01)	(0.01)	(0.02)	(0.07)	(0.01)	(0.01)	(0.05)
Data processing (reference category: non-anonymous)							
Anonymous	0.22***	0.23***	0.20***	0.10*	0.22***	0.21***	0.26***
	(0.00)	(0.01)	(0.01)	(0.05)	(0.01)	(0.01)	(0.03)
Number of vignettes	24,767	19,881	4,636	250	16,167	8,190	410
Pseudo R-squared	0.22	0.23	0.21	0.20	0.21	0.24	0.31
Log pseudolikelihood	-13,343.7	-10,611.1	-2,552.1	-138.1	-8,816.8	-4,300.1	-197.3
Wald χ2	2,441.8***	1,910.4***	522.4***	82.5***	1,511.0***	926.8***	94.5***

Note: The table reports average marginal effects for conditional logit regressions. Standard errors are in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

People with high trust are more sensitive to financial incentives than people with low trust; this holds especially for people with high trust in BigTechs and webshops. For example, when people with high trust in BigTechs are offered 20 euros this increases the likelihood of data sharing by 29 p.p., whereas the effect is only 6 p.p. for people with low trust in BigTechs. A possible explanation is that financial incentives are more likely to be perceived as suspicious by people with low trust than by people with high trust. Furthermore, people with high trust in BigTechs and people with high trust in webshops are more likely to share their data on personal preferences than people with low trust in these firms. In addition, we find that the effect of anonymity on the likelihood of data sharing is smaller for people with high trust in BigTechs than for people with low trust in BigTechs. Intuitively, people with high trust put a lower probability on the risks that arise if personal data are shared non-anonymously. Thus, trust and anonymity are to some extent substitutes.

5.4 Financial compensation needed to trigger data sharing varies a lot

For sharing health data instead of payments data one needs to pay a compensation of 9 euros per month. This indicates how less willing people are to share health data relative to payments data. In contrast, the compensation for sharing other data types is lower than for sharing payments data. People would need 27 euros per month less in case of sharing the location data of their smartphone, 9 euros less for sharing data on wealth and pensions, 57 euros less for sharing data on personal characteristics, and 62 euros less for sharing data on personal preferences (Table 8 and Figure 5). The ordering of the different types of data across gender, education, age group and income is very consistent. This ordering is also the same for people who differ with respect to their digital skills, webshop usage and social media usage (Table A.3 in Appendix A). The ordering is slightly different for people with high trust in insurers, BigTechs and/or webshops (Table A.4 in Appendix A).

The financial compensation that firms need to offer such that people share their data with them instead of with banks varies by type of firm and depends on consumer characteristics (Table 8 and Figure 6). On average, insurance firms need to offer 17 euros per month, BigTechs 32 euros and webshops 39 euros. This ordering is very consistent across gender, education, age group and income. The compensation needed is relatively high for women, old people, less educated people, and people with a low income. It is also relatively high for people with a low level of digital literacy, people who do not use social media and those who do not shop online (Table A.3). The compensation needed from a firm depends on people's trust in the firm. For example, BigTechs need to offer people with low trust in BigTechs 41 euros and people with high trust in BigTechs only 4 euros (Table A.4).

Table 8. Financial compensation needed in exchange for sharing data

In euros per month with 95% confidence intervals

		(1) All	(2) Men	(3) Women	(4) Age <45	(5) Age ≥45	(6) Low education	(7) High education	(8) Low income	(9) Middle income	(10) High income
		Compensation needed to share other type of data than payments data									
Health data	**WTA**	**8.63**	**11.19**	**3.41**	**1.57**	**13.93**	**4.48**	**13.07**	**-2.41**	**6.26**	**15.89**
	low	3.54	5.79	-8.47	-4.14	5.54	-3.10	6.63	-17.20	-3.07	8.77
	high	13.72	16.58	15.29	7.27	22.32	12.06	19.50	12.39	15.59	23.00
Location data smartphone	**WTA**	**-26.90**	**-17.11**	**-52.12**	**-12.00**	**-38.09**	**-31.23**	**-22.34**	**-34.40**	**-31.08**	**-19.88**
	low	-34.31	-23.26	-80.86	-18.76	-51.99	-43.89	-30.52	-60.48	-46.54	-27.94
	high	-19.49	-10.95	-23.38	-5.25	-24.20	-18.58	-14.16	-8.31	-15.62	-11.81
Wealth and pensions	**WTA**	**-8.56**	**-5.54**	**-16.53**	**-8.80**	**-8.35**	**-9.98**	**-6.97**	**-19.36**	**-6.19**	**-6.54**
	low	-13.49	-10.32	-30.88	-14.45	-15.88	-17.91	-12.73	-38.94	-14.82	-12.68
	high	-3.62	-0.76	-2.19	-3.15	-0.82	-2.06	-1.22	0.22	2.43	-0.40
Personal characteristics	**WTA**	**-56.88**	**-36.03**	**-111.79**	**-37.59**	**-71.64**	**-67.35**	**-45.85**	**-86.56**	**-63.33**	**-41.71**
	low	-69.85	-45.27	-169.90	-49.21	-95.41	-91.02	-58.88	-144.10	-90.64	-54.29
	high	-43.90	-26.79	-53.67	-25.98	-47.87	-43.68	-32.82	-29.01	-36.02	-29.13
Personal preferences	**WTA**	**-61.51**	**-37.19**	**-125.72**	**-44.11**	**-74.78**	**-77.02**	**-45.29**	**-99.03**	**-67.73**	**-44.44**
	low	-75.39	-46.62	-190.60	-56.81	-99.63	-103.57	-58.37	-164.58	-96.75	-57.59
	high	-47.64	-27.76	-60.84	-31.42	-49.92	-50.47	-32.22	-33.48	-38.71	-31.29
Compensation needed to share data with another firm than a bank											
Insurer	**WTA**	**17.25**	**12.98**	**28.65**	**9.30**	**23.29**	**22.47**	**12.04**	**24.66**	**18.11**	**15.10**
	low	12.32	8.55	12.00	4.75	14.41	13.42	6.98	5.79	8.49	9.28
	high	22.18	17.41	45.30	13.84	32.18	31.52	17.11	43.52	27.73	20.91
BigTech	**WTA**	**31.93**	**25.74**	**48.65**	**16.97**	**43.06**	**40.21**	**23.47**	**40.79**	**35.52**	**28.03**
	low	24.21	18.85	22.41	10.79	28.27	25.74	15.89	12.62	19.13	19.20
	high	39.65	32.63	74.89	23.16	57.86	54.67	31.05	68.96	51.90	36.86
Webshop	**WTA**	**39.25**	**31.91**	**59.21**	**20.50**	**53.66**	**50.11**	**28.21**	**47.10**	**44.58**	**33.94**
	low	30.13	23.91	27.58	13.65	35.66	32.41	19.80	14.60	24.84	23.78
	high	48.37	39.91	90.84	27.35	71.66	67.81	36.61	79.61	64.31	44.09
Compensation needed to share data non-anonymously											
Non-anonymously	**WTA**	**72.21**	**54.06**	**121.46**	**45.08**	**92.79**	**78.46**	**63.57**	**92.67**	**70.87**	**64.90**
	low	56.85	41.99	59.99	33.50	63.15	52.09	47.64	33.14	41.31	47.85
	high	87.57	66.13	182.94	56.66	122.43	104.83	79.49	152.20	100.43	81.95

Note: These are estimates of willingness-to-accept (WTA). These are based on the results of conditional logit regressions in which reward is included as a continuous variable. See Table A.2 of Appendix A for the results of these regressions.

Figure 5. Financial compensation needed for sharing other types of data than payments data
in euros per month

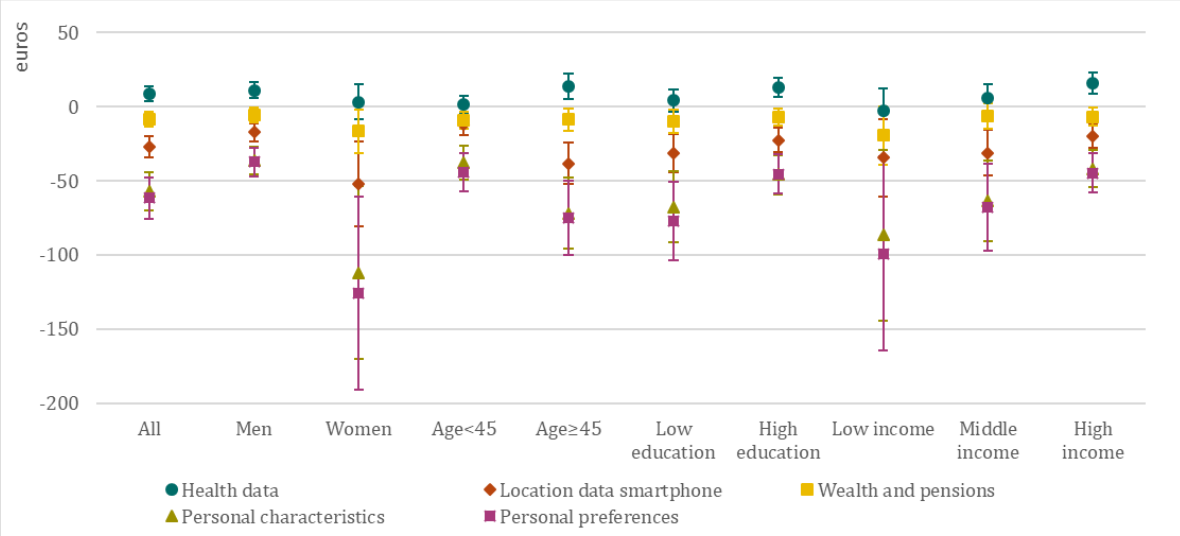
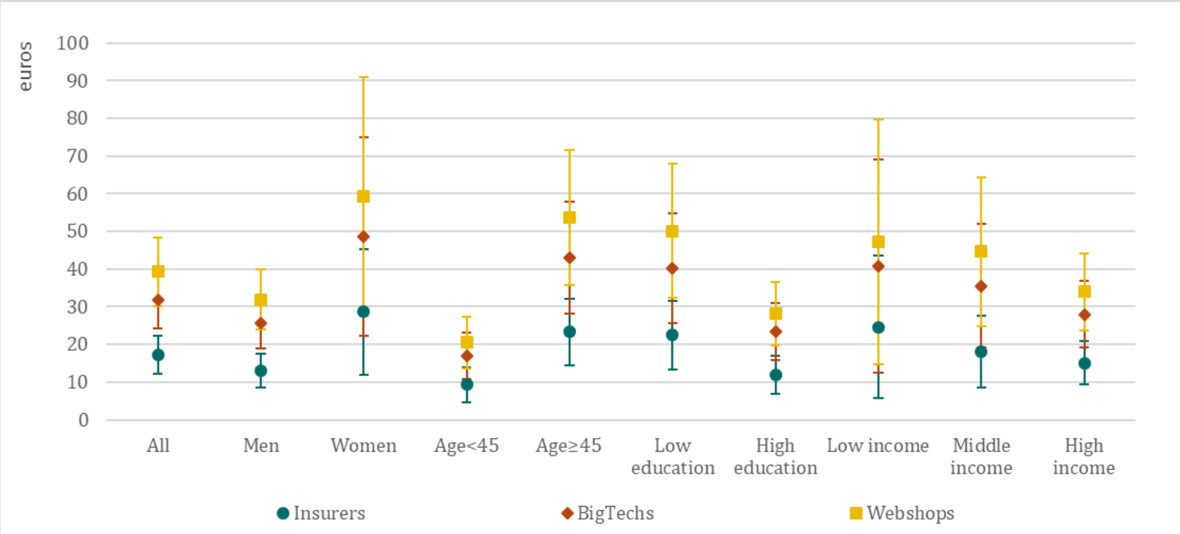
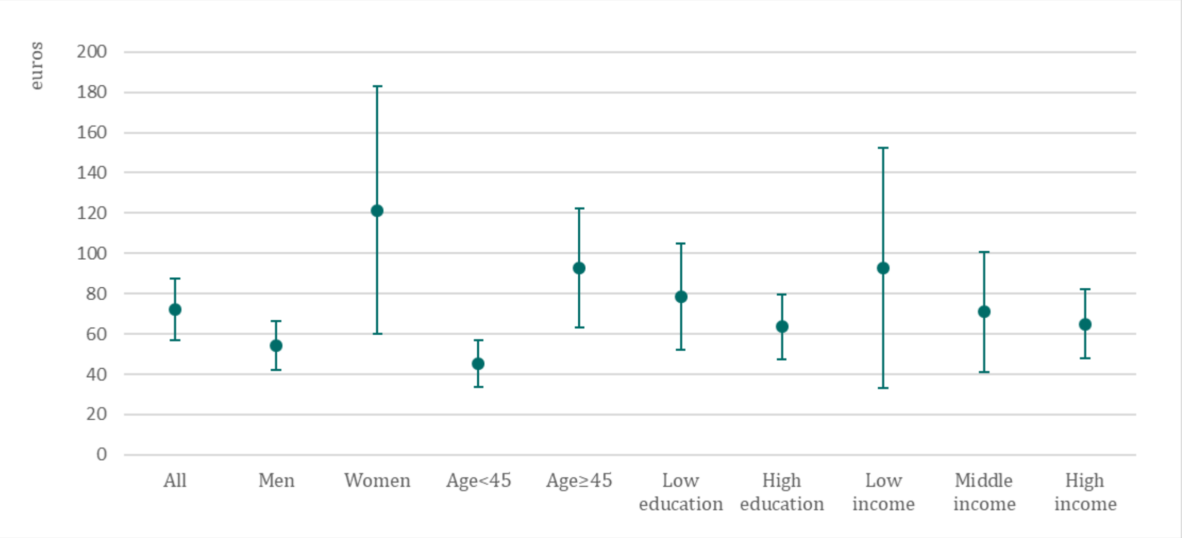


Figure 6. Financial compensation needed for sharing data with other firms than banks
in euros per month



The necessary compensation for non-anonymously sharing data is on average 72 euros per month. It is relatively high for women, old people, less educated people, and people with a low income (Table 8 and Figure 7). For example, the compensation needed is 93 euros for people aged 45 or above and 45 euros for younger people. The compensation is relatively low for people with a high level of digital literacy, who use social media a lot and/or are frequent users of webshops. For example, people who use social media more than once a day need a compensation of 56 euros for non-anonymous usage of their data, whereas people who never use social media such as Instagram, WhatsApp, Facebook, Twitter or YouTube require 132 euros (Table A.3). The compensation needed is relatively high for people with little trust in the data-using firms.

Figure 7. Financial compensation needed for using data non-anonymously
in euros per month



For example, it is 89 euros for people who distrust BigTechs and 9 euros for people who trust BigTechs (Table A.4).

For all attributes, the variance in the willingness to pay is highest for women, high age groups, people with low education, people with low income, people with a low level of digital literacy, people who do not use social media and people who do not use webshops. Within these groups there are people who can be persuaded by relatively low levels of compensation, but also people that require relatively high financial rewards to provide access to their data.

6. Concluding remarks

PSD2 has opened the possibility in the EU for third parties to access consumers' payment accounts to use their payments data to provide them with new services, and regulation is on its way to allow third parties to gain access to other data as well as part of the EU's digital markets agenda. This is also a trend in other parts of the world. Using the results of a discrete choice experiment among a representative group of Dutch consumers, we study how consumers' willingness to give firms access to their personal data depends on the type of data, the type of firm, financial incentives and anonymity.

Our results show that consumers are least likely to share data about their health, followed by data on their payments and data on wealth and pensions. These data types are less likely to be shared compared to the other three considered data types, i.e. data on the location of their smartphone, their personal characteristics and their preferences. People are especially cautious to share data when they are not used anonymously by firms, but can be linked to them. However, financial rewards can trigger data sharing.

In addition, we find that people are most hesitant to share their data with webshops and BigTechs and are also less likely to share their data with insurers than they are with banks. Trust plays a role here, with people having most trust in banks and least trust in webshops and BigTechs. This finding suggests that banks, and to a lesser extent insurers, are in a strong position to exploit the possibilities of data sharing and improve their services. Especially people aged 45 and over and high-income people prefer to share data with banks over sharing it with other firms.

The amount people need to receive to give consent varies most by the way their data are processed, followed by the type of data that is shared and varies least with the type of firm that receives the data. This suggests that if other firms turn out to be quicker and better in using customers' personal data and are able to pass on financial benefits to their (potential) customers, these firms may compete successfully with banks and get access to various types of personal data from consumers. If that happens, risks of data concentration may arise, as firms like BigTechs already have huge amounts of consumer data in their possession.

The survey also shows that vulnerable people with a low income are relatively less sensitive to financial incentives given by firms than people with a high income. This suggests that although people may be less well off financially, they may not be seduced more easily with small financial benefits than others. However, we also find that people with little education, low income and low digital skills are more likely to share their data non-anonymously with firms than others, which indicates that data sharing with third parties may endanger data privacy of vulnerable consumer segments more than of other people.

We also find that most people would rather not give third parties access to their payments data in exchange for services. Further growth in data services could be stimulated by giving consumers more control over the use of their own data and providing them with better insight which parties have access to which data. This might solve part of the confidence issue and will allow the public to reap more benefits from such data use. In addition, allowing for further data sharing should be accompanied with public campaigns with special attention to vulnerable consumers to inform them well about the possible benefits and the risks of data sharing for them.

Allowing for more data sharing may also lead to increased and new risks. It should therefore be accompanied by regulation that adequately balances different public goals such as innovation and competition on the one hand, and ensuring consumer protection, data privacy and financial stability on the other hand. This requires close co-operation between different supervisors with different mandates, both nationally and internationally. The increased risk of data concentration warrants special attention by regulators as the relevant regulatory frameworks may need to be adjusted to adequately address the risks associated with data concentration.

References

Acquisti, A., C. Taylor and L. Wagman (2016). The economics of privacy. *Journal of Economic Literature* 54(2), 442–92.

Acquisti, A., L.K. John and G. Loewenstein (2013). What is privacy worth? *The Journal of Legal Studies* 42, 249–274.

Armantier, O., Doerr, S., Frost, J., Fuster, A. and K. Shue (2021). Whom do consumers trust with their data? US survey evidence. BIS Bulletin 42, Bank for International Settlements, Basel.

Athey, S., C. Catalini and C. Tucker (2017). The digital privacy paradox: small money, small costs, small talk. NBER Working Paper 23488. NBER, Cambridge, MA.

Bansal, G., F.M. Zahedi and D. Gefen (2016). Do context and personality matter? Trust and privacy concerns in disclosing private information online. *Information & Management* 53, 1–21.

Benndorf, V. and H-T. Normann (2018). The willingness to sell personal data. *The Scandinavian Journal of Economics* 120(4), 1260–1278.

Bijlsma, M, C. van der Cruijsen and N. Jonker (2020). Consumer propensity to adopt PSD2 services: trust for sale? DNB Working Paper 671. DNB, Amsterdam.

Carlsson. F. and P. Martinsson (2003). Design techniques for stated preference methods in health economics*. Health Economics* 12, 281–294.

Carrière-Swallow, Y., V. Haksar and M. Patnam (2021). India's approach to open banking: some implications for financial inclusion. IMF Working paper WP21/52. IMF, Washington D.C..

Choi, J.P., D.-S. Jeon and B.-C. Kim (2019). Privacy and personal data collection with information externalities. *Journal of Public Economics* 173, 113–124.Economic Commission (2020). A European strategy for data. European Commission, Brussels.

Garratt R. and M. van Oordt (2021). Privacy as a Public Good: A Case for Electronic Cash. *Journal of Political Economy*. 129(7), 2157-2180.

Goldfarb, A. and C. Tucker (2012). Shifts in privacy concerns. *American Economic Review* 102, 349–53.

Hann, I.-H., K.-L. Hui, S.-Y.T. Lee, and I.P. Png, (2007). Overcoming online information privacy concerns: an information-processing theory approach. *Journal of Management Information Systems* 24, 13–42.

Hole. A.R. (2016). Creating efficient designs for discrete choice experiments. Presentation given at the Nordic and Baltic Stata Users Group Meeting. September 2016.

McKinsey & Company (2021) Financial data unbound: The value of open data for individuals and institutions. Discussion paper. McKinsey Global institute.

OECD (2019). Enhancing Access to and Sharing of Data. Reconciling Risks and benefits for Data Re-use across Societies. OECD Publishing, Paris.

Prince. J. and S. Wallsten (2020). How much is privacy worth around the world and across platforms? Technology Policy Institute. Washington D.C.

Regner, T. and G. Riener (2017). Privacy is precious: on the attempt to lift anonymity on the internet to increase revenue. *Journal of Economics & Management Strategy* 26(1), 318–336.

Van der Cruijsen, C. and F. van der Horst (2019). Cash or card? Unravelling the role of socio-psychological factors. *De Economist* 167(2), 145–175.

Van der Cruijsen, C. (2020). Payments data: do consumers want banks to keep them in a safe or turn them into gold? *Applied Economics* 52(6), 609–622.

Van der Cruijsen, C., J. de Haan and N. Jonker (2021). Has the COVID-19 pandemic affected public trust? Evidence for the US and the Netherlands. DNB Working Paper 723. DNB, Amsterdam.

Teppa. F. and C. Vis. (2012). The CentERpanel and the DNB Household Survey: methodological aspects. DNB Occasional Study 10(4). DNB, Amsterdam.

Zwerina, K., Huber, J. and W.F. Kuhfeld. (1996). A general method for constructing efficient choice designs. Working Paper. Fuqua School of Business, Duke University, Durham N.C.

Appendix A: Additional results

Table A.1. Regression results: detailed age groups

	(1) All	(2) Age ≤24	(3) Age 25-34	(4) Age 35-44	(5) Age 45-54	(6) Age 55-64	(7) Age ≥65
Data type (reference category: payments data)							
Health data	-0.03***	0.03	0.01	-0.03	-0.04***	-0.05***	-0.02
	(0.01)	(0.04)	(0.03)	(0.02)	(0.01)	(0.02)	(0.01)
Location data smartphone	0.09***	0.01	0.07**	0.07***	0.08***	0.11***	0.09***
	(0.01)	(0.04)	(0.03)	(0.02)	(0.01)	(0.02)	(0.01)
Wealth and pensions	0.03***	0.05	0.03	0.06***	0.03**	0.02	0.02*
	(0.01)	(0.03)	(0.02)	(0.02)	(0.01)	(0.02)	(0.01)
Personal characteristics	0.17***	0.17***	0.20***	0.17***	0.16***	0.18***	0.17***
	(0.01)	(0.03)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
Personal preferences	0.18***	0.25***	0.19***	0.20***	0.19***	0.18***	0.17***
	(0.01)	(0.03)	(0.02)	(0.02)	(0.01)	(0.02)	(0.01)
Firm (reference category: bank)							
Insurer	-0.05***	-0.04*	-0.04**	-0.04***	-0.04***	-0.06***	-0.06***
	(0.01)	(0.03)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
BigTech	-0.09***	-0.07**	-0.08***	-0.09***	-0.09***	-0.08***	-0.11***
	(0.01)	(0.03)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
Webshop	-0.12***	-0.15***	-0.08***	-0.09***	-0.10***	-0.09***	-0.15***
	(0.01)	(0.04)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
Financial compensation (reference category: no compensation)							
2 euros	0.02***	0.06**	0.04**	0.04**	0.03***	0.00	0.02*
	(0.01)	(0.03)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
5 euros	0.03***	0.05	0.02	0.05***	0.07***	0.01	0.01
	(0.01)	(0.03)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
10 euros	0.05***	0.09***	0.08***	0.07***	0.08***	0.05***	0.02
	(0.01)	(0.03)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
20 euros	0.07***	0.14***	0.11***	0.10***	0.12***	0.06***	0.01
	(0.01)	(0.03)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
Data processing (reference category: non-anonymous)							
Anonymous	0.22***	0.21***	0.21***	0.23***	0.23***	0.24***	0.20***
	(0.00)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Number of vignettes	24,767	900	2,193	3,392	7,581	4,940	9,153
Pseudo R-squared	0.22	0.26	0.24	0.26	0.27	0.26	0.18
Log pseudolikelihood	-13343.7	-458.7	-1160.3	-1731.1	-3850.4	-2538.1	-5199.2
Wald $\chi 2$	2441.8***	231.8***	266.6***	464.2***	948.7***	583.0***	749.3***

Note: The table reports average marginal effects for conditional logit regressions. Standard errors are in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table A.2. Regression results: demographic groups and linear reward variable

	(1) All	(2) Men	(3) Women	(4) Age <45	(5) Age ≥45	(6) Low education	(7) High education	(8) Low income	(9) Middle income	(10) High income
Data type (reference category: payments data)										
Health data	-0.029***	-0.052***	-0.007	-0.009	-0.037***	-0.012	-0.062***	0.006	-0.019	-0.071***
	(0.008)	(0.012)	(0.012)	(0.016)	(0.010)	(0.010)	(0.014)	(0.018)	(0.014)	(0.014)
Location data smartphone	0.088***	0.076***	0.100***	0.065***	0.096***	0.083***	0.098***	0.079***	0.092***	0.084***
	(0.008)	(0.011)	(0.011)	(0.016)	(0.009)	(0.010)	(0.013)	(0.017)	(0.014)	(0.012)
Wealth and pensions	0.029***	0.025**	0.033***	0.048***	0.022**	0.027***	0.032**	0.045***	0.019	0.029**
	(0.008)	(0.011)	(0.011)	(0.014)	(0.009)	(0.010)	(0.012)	(0.017)	(0.013)	(0.013)
Personal characteristics	0.175***	0.153***	0.199***	0.185***	0.171***	0.170***	0.186***	0.185***	0.177***	0.167***
	(0.007)	(0.010)	(0.010)	(0.013)	(0.008)	(0.009)	(0.010)	(0.015)	(0.012)	(0.011)
Personal preferences	0.187***	0.157***	0.219***	0.210***	0.178***	0.191***	0.184***	0.207***	0.188***	0.176***
	(0.007)	(0.010)	(0.010)	(0.012)	(0.008)	(0.009)	(0.011)	(0.015)	(0.012)	(0.011)
Firm (reference category: bank)										
Insurer	-0.051***	-0.052***	-0.049***	-0.043***	-0.053***	-0.054***	-0.047***	-0.050***	-0.048***	-0.057***
	(0.006)	(0.008)	(0.008)	(0.010)	(0.007)	(0.007)	(0.009)	(0.012)	(0.009)	(0.009)
BigTech	-0.097***	-0.108***	-0.085***	-0.082***	-0.101***	-0.099***	-0.094***	-0.085***	-0.098***	-0.110***
	(0.006)	(0.009)	(0.009)	(0.012)	(0.007)	(0.008)	(0.010)	(0.013)	(0.011)	(0.010)
Webshop	-0.120***	-0.136***	-0.104***	-0.100***	-0.128***	-0.125***	-0.115***	-0.099***	-0.124***	-0.135***
	(0.007)	(0.010)	(0.009)	(0.013)	(0.008)	(0.008)	(0.011)	(0.014)	(0.011)	(0.011)
Financial compensation (in euros per month)										
Financial compensation	0.003***	0.004***	0.002***	0.005***	0.002***	0.003***	0.004***	0.002***	0.003***	0.004***
	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)
Data processing (reference category: non-anonymous)										
Anonymous	0.223***	0.230***	0.215***	0.222***	0.222***	0.197***	0.261***	0.197***	0.198***	0.260***
	(0.005)	(0.006)	(0.007)	(0.008)	(0.005)	(0.006)	(0.007)	(0.010)	(0.008)	(0.007)
Number of vignettes	24,767	12,799	11,968	6,485	18,282	15,377	9,370	5,376	8,703	9,348
Pseudo R-squared	0.22	0.23	0.22	0.25	0.22	0.18	0.31	0.19	0.19	0.29
Log pseudolikelihood	-13352.7	-6827.5	-6489.0	-3375.0	-9947.1	-8720.1	-4500.9	-3035.6	-4912.8	-4597.7
Wald $\chi 2$	2445.4***	1402.1***	1070.2***	851.2***	1686.5***	1321.3***	1364.0***	432.8***	793.1***	1239.1***

Note: The table reports average marginal effects for conditional logit regressions. Standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.3. Financial compensation needed in exchange for sharing data: digital skills, webshop usage and social media usage
In euros per month with 95% confidence intervals

		(1) All	(2) Low digital literacy	(3) High digital literacy	(4) Webshop usage: no	(5) Webshop usage: low	(6) Webshop usage: high	(7) Social media usage: no	(8) Social media usage: low	(9) Social media usage: high
Compensation needed to share other type of data than payments data										
Health data	**WTA**	**8.63**	**9.67**	**8.29**	**-6.33**	**9.35**	**12.98**	**15.68**	**13.97**	**7.29**
	low	3.54	-12.01	3.60	-37.62	4.16	1.05	-26.67	-5.93	2.47
	high	13.72	31.35	12.98	24.96	14.55	24.90	58.03	33.87	12.11
Location data smartphone	**WTA**	**-26.90**	**-64.11**	**-20.41**	**-80.31**	**-21.85**	**-23.07**	**-43.59**	**-56.11**	**-20.35**
	low	-34.31	-120.77	-26.55	-194.63	-28.44	-40.30	-127.03	-99.30	-26.65
	high	-19.49	-7.44	-14.27	34.01	-15.25	-5.84	39.84	-12.92	-14.04
Wealth and pensions	**WTA**	**-8.56**	**-18.28**	**-6.68**	**-28.44**	**-5.66**	**-11.27**	**-1.45**	**-13.19**	**-7.82**
	low	-13.49	-42.71	-11.18	-79.10	-10.34	-22.86	-36.91	-31.53	-12.60
	high	-3.62	6.14	-2.17	22.22	-0.98	0.31	34.01	5.15	-3.04
Personal characteristics	**WTA**	**-56.88**	**-130.44**	**-44.15**	**-165.62**	**-45.93**	**-50.59**	**-112.46**	**-97.79**	**-45.90**
	low	-69.85	-241.83	-54.22	-398.12	-56.70	-78.79	-310.66	-171.15	-56.39
	high	-43.90	-19.05	-34.09	66.88	-35.16	-22.39	85.73	-24.42	-35.41
Personal preferences	**WTA**	**-61.51**	**-138.95**	**-48.39**	**-185.76**	**-49.33**	**-53.89**	**-125.88**	**-108.50**	**-49.02**
	low	-75.39	-257.19	-59.24	-446.70	-60.72	-82.39	-351.56	-189.39	-59.97
	high	-47.64	-20.71	-37.55	75.17	-37.93	-25.39	99.80	-27.60	-38.07
Compensation needed to share data with another firm than a bank										
Insurer	**WTA**	**17.25**	**40.10**	**13.26**	**75.72**	**11.19**	**15.43**	**59.12**	**26.51**	**13.06**
	low	12.32	3.84	9.19	-31.70	7.14	4.71	-46.55	4.31	8.89
	high	22.18	76.35	17.33	183.13	15.24	26.15	164.80	48.71	17.23
BigTech	**WTA**	**31.93**	**67.92**	**25.49**	**116.20**	**23.70**	**27.27**	**109.77**	**42.96**	**25.38**
	low	24.21	8.99	19.21	-48.77	17.51	12.33	-83.83	9.14	19.08
	high	39.65	126.86	31.77	281.17	29.88	42.21	303.36	76.77	31.68
Webshop	**WTA**	**39.25**	**92.18**	**29.96**	**165.70**	**26.87**	**31.81**	**125.26**	**70.74**	**28.28**
	low	30.13	12.67	22.96	-67.63	20.19	14.12	-93.87	17.66	21.44
	high	48.37	171.70	36.96	399.03	33.54	49.50	344.39	123.81	35.13
Compensation needed to share data non-anonymously										
Non-anonymously	**WTA**	**72.21**	**156.23**	**57.48**	**213.79**	**58.94**	**59.37**	**132.33**	**136.45**	**56.29**
	low	56.85	24.22	45.80	-84.74	46.53	30.27	-100.08	36.84	67.78
	high	87.57	288.23	69.15	512.32	71.35	88.47	364.74	236.06	44.80

Note: These are estimates of willingness-to-accept (WTA). These are based on the results of conditional logit regressions in which reward is included as a continuous variable.

Table A.4. Financial compensation needed in exchange for sharing data: trust levels

In euros per month with 95% confidence intervals

		(1) All	(2) Own main bank: low trust	(3) Own main bank: high trust	(4) Insurers: low trust	(5) Insurers: high trust	(6) BigTechs: low trust	(7) BigTechs: high trust	(8) Webshops: low trust	(9) Webshops: high trust
		Compensation needed to share other type of data than payments data								
Health data	**WTA**	**8.63**	**12.35**	**6.23**	**10.67**	**3.86**	**11.50**	**7.05**	**11.06**	**9.67**
	low	3.54	-4.98	-0.45	3.06	-15.37	4.45	-3.30	3.27	-7.72
	high	13.72	29.67	12.91	18.28	23.09	18.56	17.40	18.85	27.06
Location data smartphone	**WTA**	**-26.90**	**-5.03**	**-20.49**	**-26.88**	**-22.17**	**-35.43**	**-1.33**	**-32.42**	**-4.82**
	low	-34.31	-20.27	-29.36	-37.62	-49.02	-47.07	-9.16	-44.76	-22.60
	high	-19.49	10.21	-11.62	-16.13	4.67	-23.79	6.50	-20.08	12.95
Wealth and pensions	**WTA**	**-8.56**	**-3.06**	**-9.22**	**-5.66**	**-17.17**	**-10.50**	**1.48**	**-6.84**	**10.33**
	low	-13.49	-16.57	-16.14	-12.49	-41.92	-17.22	-6.14	-14.08	-6.42
	high	-3.62	10.44	-2.30	1.16	7.57	-3.78	9.11	0.39	27.08
Personal characteristics	**WTA**	**-56.88**	**-25.13**	**-50.79**	**-49.57**	**-54.24**	**-68.45**	**-14.80**	**-64.85**	**-24.24**
	low	-69.85	-48.35	-66.96	-66.75	-102.24	-88.49	-27.90	-86.53	-49.74
	high	-43.90	-1.91	-34.63	-32.38	-6.24	-48.40	-1.70	-43.16	1.25
Personal preferences	**WTA**	**-61.51**	**-33.61**	**-55.71**	**-53.48**	**-44.30**	**-73.32**	**-11.27**	**-70.79**	**-17.00**
	low	-75.39	-62.21	-73.05	-71.69	-89.90	-94.56	-25.76	-94.13	-42.15
	high	-47.64	-5.01	-38.36	-35.26	1.29	-52.08	3.21	-47.45	8.15
Compensation needed to share data with another firm than a bank										
Insurer	**WTA**	**17.25**	**14.65**	**17.72**	**18.72**	**5.82**	**20.82**	**5.91**	**22.34**	**7.90**
	low	12.32	0.64	10.90	11.20	-7.70	13.61	-2.06	14.03	-2.73
	high	22.18	28.67	24.54	26.24	19.34	28.03	13.87	30.65	18.53
BigTech	**WTA**	**31.93**	**17.46**	**31.37**	**30.25**	**26.17**	**41.01**	**3.99**	**40.09**	**21.95**
	low	24.21	-0.42	20.98	19.23	2.05	28.61	-3.81	26.42	-0.88
	high	39.65	35.35	41.77	41.28	50.30	53.40	11.80	53.76	44.78
Webshop	**WTA**	**39.25**	**19.32**	**40.18**	**37.20**	**40.57**	**48.34**	**10.67**	**51.61**	**25.52**
	low	30.13	2.41	27.44	24.14	6.83	33.99	0.40	34.47	2.92
	high	48.37	36.23	52.92	50.26	74.31	62.69	20.94	68.75	48.11
Compensation needed to share data non-anonymously										
Non-anonymously	**WTA**	**72.21**	**62.75**	**58.07**	**73.99**	**55.16**	**88.87**	**8.54**	**85.30**	**37.60**
	low	56.85	18.13	41.13	50.56	9.48	64.17	-1.11	58.38	6.13
	high	87.57	107.36	75.01	97.41	100.84	113.56	18.19	112.22	69.08

Note: These are estimates of willingness-to-accept (WTA). These are based on the results of conditional logit regressions in which reward is included as a continuous variable.

DeNederlandscheBank
EUROSYSTEEM

De Nederlandsche Bank N.V.
Postbus 98, 1000 AB Amsterdam
020 524 91 11
dnb.nl