Pre-Analysis Plan: *Preferences for the use and regulation of algorithms in the context of self-driving cars and algorithmic decision-making*

Sebastian Hemesath (University of Oldenburg)
Markus Tepe (University of Oldenburg)

02.12.2020

Project overview and introduction

The increased use of algorithmic decision-making (ADM) systems in many domains of public life raises questions regarding the scope and extent towards which societies are willing to transfer human decision-making to ADM systems, as well as the ways those novel technologies should be governed politically. As part of a bigger research project on the social embeddedness of autonomous cyber-physical systems, this study explores attitudes and preferences of individuals regarding the use and regulation of algorithms and algorithmic systems, using original data from representative online surveys administered in Germany, Japan and the United States. The survey will consist of three separate experiments, capturing three different approaches to this issue: First, to analyse citizens' preferences regarding the regulation of self-driving cars, one of most widely known use-cases of algorithmic decision-making for individuals, we will conduct an experimental conjoint analysis on regulatory preferences, to understand which design-features of self-driving car regulation promote mass support for regulation. Second, as policies are subject to public deliberations, and public support for policy-implementation often succeeds or dies according to the salience of the dominant policy frames, we evaluate the effect of different policy frames on support for self-driving car approval, by conducting a framing experiment. Last, to evaluate the circumstances under which individuals accept the use of ADM over human decision-making in general, we will conduct a vignette experiment on the acceptance of ADM in the context of hypothetical hiring scenarios.

Reasoning for sample selection and technical implementation

The survey will be conducted in three countries (Germany, Japan and the United States), where we aim to draw a representative sample from the population of the voting-age citizens in each nation. The choice for those countries is based on three criterions:

1.) The importance of the automotive industry for the overall economy of the individual country, in order to ensure that the regulation of self-driving cars possesses some salience for individuals, and it is more likely that individuals already are in possession of a broad mental model of self-driving cars.
2.) Cultural variations between those countries, as we expect cultural differences regarding the use and associated risk with novel technologies
3.) Since the goal of the study is to analyse citizens preferences and attitudes towards political regulation, the classification of the individual country as a liberal democracy, with a functioning civic culture.

The sample size for the study is set at 3.600 participants, with the individual samples of each country aimed to estimate approximately 1.200 respondents. The sample size was a trade-off between budgetary limitations and the expected statistical power, estimated by power analyses. The responses will be collected using market-research firm "respondi", in order to collect a representative sample of the target populations. To ensure a representative distribution of the sample, we implement uncrossed quotas for age (4 groups), education (3 groups), gender (2 groups) and region (state-level or census regions), which are based on the demographic distributions retrieved from the individual national census statistics (Table 1). Based on the approximated response-time of 10-15 minutes, respondents that complete the survey will receive a monetary incentive of 0.60 EUR. The survey is programmed and hosted on the online survey platform "qualtrics". For each country, we translated the survey into the national language (German, English, Japanese).

Conjoint Experiment

Description and theoretical background

Effective regulation of self-driving cars requires broad public support. Although policy-makers and practitioners agree upon the growing need to regulate the development of autonomous cyber physical systems (ACPS) (e.g. European Commission 2019), and the importance of regulation that is consistent with citizens' moral believes and societies' legal standards, there is little systematic research about which type of regulation citizens prefer and whether the public is sensitive to the specific features of different ACPS regulation architectures. This is of relevance in at least two ways: First, for public policy to be legitimate, it needs to align with societal preferences (Savulescu et.al 2019) and consider existing preference trade-offs in society (Montpetit 2008, Wallner 2008). Designing effective regulatory policy requires the balancing of societal and stakeholder preferences with effective tools and instruments to achieve the regulatory goals (Majone 1975). Second, building on insights from theories on socio-technological systems and technological transitions (e.g. Geels 2007, Rip and Kemp 1998), the regulatory landscape influences the development and success of technological transitions, creating path dependencies by shaping its design, trajectory and characteristics (Smith et.al 2005). Regulation, depending on scope, point in time, and the specific regulatory measures applied, can stall or enable innovation (Weiner 2004). The closer the regulatory framework is to the preferences of society, the more likely it will thus become that society will trust and ultimately adopt the technology (Liu et.al 2018). As regulation that is in line with societal preferences can mitigate the anticipated risks and undesired consequences of technologies (Weiner 2004, Haines 2011, Reins 2019). Corresponding evidence from the field of technological acceptance shows that trust is a necessary link between individuals' beliefs towards automated technologies and their subsequent intention to use them (Pavlou 2003, Choi and Ji 2015). However, much of the empirical research on societal acceptance of AVs is limited to single regulatory dimensions, for example the empirical investigation of the trolley dilemma in moral conflict situations (Bonnefon et al. 2016, Shariff et al. 2017), or acceptance of autonomous vehicles in the context of future AV adoption (see Gkartzonikas and Gkritza 2019 for an extensive review). (Regulatory) Policy preferences on the other hand, are always a multidimensional aggregation of issue preferences where, depending on the issue and dimension of regulation, different attributes of the policy package are more salient for individuals than others (Bechtel et.al 2012). Limiting regulatory preferences to single dimensions would thus be unrealistic and ignore the preference trade-offs that are inherent to

regulation. For this reason, we designed a conjoint experiment to analyse the effect of different regulatory attributes on public support for self-driving car regulation. Conjoint experiments allow disentangling the preferences towards different elements of policy packages, in order to measure the causal effect of different policy features on support for regulation.

Experimental Design

This part of the study uses a fully randomized paired-conjoint experiment (Hainmueller, Hopkins and Yamamoto 2014, Hainmueller 2015), in which respondents are faced with two hypothetical proposals for self-driving car regulation and have to choose one. The specific features (levels) of the proposals vary randomly across a set of regulatory attributes, whereas the attributes represent three broad regulatory dimensions:

1. The technical regulation of self-driving cars (regulation to mitigate technical risk)
2. The regulation of interaction with and between users of the technology (risk of behavioural interaction)
3. The loss of human agency by the transfer of human decision-making to autonomous technologies (broader societal risk)

For each regulatory dimension, we defined two regulatory attributes with different levels that measure the most important preference trade-offs in each dimension, resulting in six comprehensive and non-overlapping attributes (Tab.1). Before conducting the experiment, we introduce the regulatory attributes to the respondents in detail (Tab.2). The respondents are then tasked to select their preferred proposal for regulation from the choice of two proposals (Fig.1), while the levels of the attributes change for each task within subject, and the order of the attributes varies between subjects, in order to prevent order-effects. In total, the experiment will be repeated six times for each participant. For the dependent variable, we ask respondents to select their preferred design out of the two present options (binary choice)*:*

- Please carefully review the two legislative proposals. Which proposal do you prefer?[Proposal A; Proposal B]

To identify the determinants of regulatory preferences, we will additionally collect responses for three covariates, which will be used for the subgroup analysis (Table 3). Based on theoretical considerations from the literature, we expect regulatory preferences for self-driving cars to be determined by political ideology, cultural worldviews and risk perception, as well as individuals' attitudes towards technology.

Analysis

The conjoint experiment will be analysed using standard methods, estimating the marginal means (MM) for each attribute level, that is, the marginal probability that the individual profile is selected, when the specific attribute level is present (binary dependent variable). The calculation of marginal means is selected over the dominant approach of estimating average marginal component effects (AMCE), as research (Leeper 2018) suggests that marginal means are a more valid descriptive estimate when conducting subgroup analyses, since MM in contrast to AMCE don't rely on reference categories. Following the recommendations from Hainmueller et al. (2014), we estimate the MM by using linear probability models and regress the dependent variable on dummy variables for each attribute level. Besides the baseline effects for each attribute level, we will also estimate the interaction effect between different attributes. To estimate these effects we use the "cregg" package in R with clustered standard errors by

respondents, to account for correlations within responses from a given respondent. For the subgroup analysis for heterogeneous effects, we will transform the covariate indexes into categorical factors and group the baseline results from the conjoint analysis by the different covariate groups (e.g. technophile/technophobe). While the applied quotas of our design aim to account for differences in inclusion probabilities, sampling errors and possible non-responses might still occur, therefore we will use the census data to create post-stratification weights in order to adjust the sample. We will report both the weighted and unweighted results. For testing the robustness of our results, we will use a series of diagnostic tests. To test the methodological assumption of conjoint experiments that there are no carryover effects between individual responses, we will estimate the MM separately for each of the six tasks per respondent. Second, we will test whether profile order effects regarding the position of the individual regulatory proposal exist, that is whether each profile is on the left or right of the screen. Third, even though we randomized the attribute order between respondents, we will additionally test for order effects resulting from the profile order, by estimating row-specific MM. Since the survey is programmed to be used both on mobile devices as well as on desktop computers, and while we advise respondents on mobile devices to hold their device horizontally in order to display the conjoint tables correctly, we can't rule out the possibility that the screen size might nevertheless have an influence on the respondents' choices. Therefore, we will separately analyse the results from mobile and non-mobile respondents to check whether any significant differences exist.

Framing Experiment

Description and theoretical background

The design of legislation for novel technologies cuts two ways: on the one hand, regulatory measures aim at generating predictability and common standards in the face of uncertainty, in order to counteract undesired consequences such as possible safety hazards and health risks. On the other hand, those measures, depending on scope, point in time, and the specific measures applied, can stall technological innovation and hinder economic success. This trade-off in policy-making becomes apparent for the approval of self-driving cars for testing: policy-makers and administrative bodies have to find the right balance between safety demands required to test self-driving vehicles on public roads, and the pressing industrial need to test the vehicles under real-life conditions on public roads, to develop effective algorithms. This study explores the salience and preferred cost-benefit distribution of this major trade-off between safety and industrial interests among citizens of automotive nations, and how different policy frames effect public support for the approval of self-driving cars. The presence of *competitive* frames in public debate has been found to affect individuals' ability to form preferences for policies (Chong and Druckman 2007, Sniderman and Thierault 2004). To analyse public support for the approval of self-driving cars, it is therefore essential to analyse the effect of these policy frames on societal acceptance, and the associated salience of the different dimensions. Given that both frames rarely exist on their own, and evidence suggests the importance of *counter-frames (*Chong and Druckman 2011) for the successful change of public opinion, we not only test for the baseline effect of both economy and safety frames, but likewise the effect of the interaction with the respective counter-frames. That is, the introduction of the corresponding cost of the other frame towards the gain frame of the baseline condition. Since recent public opinion polls indicate individuals to be highly concerned with the possible safety risks of self-driving cars, we hypothesize that a safety frame will lead to higher public acceptance of the approval process,

than an economic frame. Nevertheless, we expect that the presence of a cost trade-off will reduce agreement with the statement for both frames, as the economic importance of this novel technology will still possess some salience in those nations. Evidence further found the sensitivity of framing effects to vary between individuals, where some individuals are more sensitive to framing effects than others. A large body of literature for example found the effect of frames on trade-policy preferences to be determined by the level of education and skills of individuals, or more generally the salience of specific frames or issues for individuals. Therefore, we hypothesise that the effect of issue frames for self-driving car regulation is similar to policy preferences in general, determined by the political ideology of individuals, cultural worldviews and associated risk perceptions, as well as the technophile beliefs of individuals (Table 4).

Experimental Design

For this experiment, respondents will be randomly assigned to one of four treatment groups in a between-subject design. The assignment to each group will occur with equal probability; therefore, approximately 275 individuals will be assigned to each group per sample. The treatment conditions represent the two baseline frames for safety and economic gain, as well as the two interactions between baseline frame and counter-frame. In each treatment group, respondents will then be asked to state their agreement with a framed statement on the approval of self-driving cars (see below). The stated agreement on a 11-point scale will serve as the outcome measure for the analysis. Below are the statements for the different treatment groups:

"Self-driving cars use machine learning. This means the earlier a car manufacturer succeeds in testing its self-driving cars in public traffic, the earlier this manufacturer can offer a fully developed product in series production. *On a scale from 0 (fully disagree) to 10 (fully agree), how much do you agree with the following statement?"*	
Economic gain baseline	*To strengthen the domestic automotive industry, self-driving cars should be approved quick and non-bureaucratic.*
Safety gain baseline	*The safety of other traffic participants must always come first when approving self-driving cars.*
Economic gain with safety cost	*To strengthen the domestic automotive industry, self-driving cars should be approved quick and non-bureaucratic, even if this puts the safety of other traffic participants at risk.*
Safety gain with economic cost	*The safety of other traffic participants must always come first when approving self-driving cars, even if this puts the domestic automotive industry at a disadvantage.*

Analysis

To test our hypotheses, and analyse the effect of the different frames on support for regulation, we will cluster the treatments into two factors, kind of policy frame and the presence of a trade-off, and regress the two factors on the respective outcomes using OLS regression

		Trade-off	
		Present	Not-present
Frame	Economy	y_{iq}	y_{iq}
	Safety	y_{iq}	y_{iq}

The linear model will estimate the following equation, where y_{iq} is the individual agreement with the statement, *Frame* and *Trade-off* are the two treatment factors, $\vartheta_i X$ is a vector of control variables, and ε the error term:

$$y_{iq} = \beta_0 + \beta_1 Frame + \beta_2 Tradeoff + \delta(Frame\ X\ Tradeoff) + \vartheta_i X + \varepsilon$$

We will estimate the linear model without covariates, with demographic covariates, as well as with a more extensive set of further explanatory variables, that is the three explanatory concepts we expect to determine differences in policy preferences among individuals and as we hypothesize also their sensitivity to different issue frames (political ideology, cultural worldviews and level of technophilia). While the applied quotas to our samples aim to account for differences in inclusion probabilities, sampling errors and possible non-responses might still occur, therefore we will use the census data to create post-stratification weights in order to adjust the sample. We will report both the weighted and unweighted results.

Vignette experiment

Description and theoretical background

While the first two parts of this study focus on preferences for regulation of self-driving cars, with the third experiment we are interested in understanding the circumstances under which individuals accept the use of ADM over human decision-making in general, and whether the context and the way how ADM is applied, influence public acceptance of the technology. To prevent priming effects from the previous experiments, and since analysing acceptance of ADM in a use-case that is not yet in existence would introduce issues of causality, we chose to apply another use-case for ADM to this experiment, and analyse acceptance in the context of hiring decisions. To use hiring scenarios is beneficial in at least three ways: First, the use of ADM in hiring decisions is already commonplace in many organizations and it is highly likely individuals are at least partially familiar with this use-case. Second, similar to the use of ADM in the context of self-driving cars, the decision-making in this case likewise carries a potential personal risk for individuals (health risks and personal safety in the case of self-driving cars vs. economic risk of not finding a job in the case of hiring decisions). Third, a large body of literature already exists on the impact of ADM in the context of hiring decisions, which allows replicating existing evidence in a comparative context. Theoretically, we start from the assumption that the acceptance of a novel technology occurs, if the benefit of adopting the technology outweighs the associated risks (extended TAM). From the literature, we identified two dimensions that could, in the sense of push and pull factors, make individuals then value the benefits of adopting the technology higher than remaining at the status quo: The first dimension refers to the utility, that is gains in efficiency algorithmic decision-making infers. The second to the perceived fairness of ADM which potentially outweighs a perceived bias of human decision-making. Therefore, we expect perceived efficiency and perceived fairness, as well as their interaction with each other, to be driving factors behind the acceptance of ADM in hiring scenarios. We further expect that differences in the perceived efficiency and fairness of ADM can be explained by individual characteristics of the respondents. On the one hand, we expect general risk perceptions, as well as attitudes to technology to influence individuals' acceptance of ADM, on the other, we hypothesize that the perceived fairness is largely determined by individuals' experiences with discrimination.

Experimental Design

To test the effect of different applications of ADM on perceived fairness, effectiveness and acceptability, we designed a between-subjects vignette experiment of hypothetical hiring scenarios and asked respondents to evaluate how fair, effective and acceptable they think this form of hiring process is. The vignettes were designed as a descriptive text, in which we embedded a 2 (Publicness: Public vs. Private) x 2 (Privacy: limited vs. extensive) x 2 (Accountability: human aided by algorithm vs. ADM without human input) factorial design (Auspurg and Hinz 2014) alongside three contextual factors that relate to the hypothetical use of ADM in hiring scenarios. The design of the descriptive text is based on a design recently used by PEW research (Smith 2018). The first factor concerns the question by whom ADM is used, and whether there are differences in acceptance between its use by public and private organizations. Building on the theory of representative bureaucracy, we would expect that individuals show more scepticism towards the use of ADM by public organizations than by private companies, thereby decreasing the perceived fairness when applied in the context of public administration. The second factor combines recent debates on transparency and privacy that ADM invokes and manipulates the data pool of the algorithm: In contrast to human decision-making, ADM offers the possibility to base the decision on a larger data pool, data-points which might also be external to the specific decision scenario itself. For the case of hiring scenarios this could be the use of additional personal data about the applicant (e.g. available data on the internet), that goes beyond the content of the application itself. While the use of such extensive data could increase the efficiency of ADM, it simultaneously raises questions of privacy, which could negatively effect the perceived fairness of such decision. The third factor concerns the decision-making power of ADM systems, and whether human decision-making is supported, or substituted by ADM. Recent evidence found trust in algorithms to depend significantly on the ability of human-subjects or decision-makers to modify the algorithms judgement (Burton et.al 2020). Applied to the hiring scenario, this means, whether the algorithmic decision is used to inform human decision-making, or the decision is made without human involvement. Below we will introduce the descriptive texts, as well as the different factorial levels we employ in the set of vignettes:

> *"In an effort to improve the hiring process, some* ***[Publicness: (1) public authorities (2) companies]*** *are now using computer programs to screen resumes.*
>
> *The computer program assigns each candidate an automated score based on* ***[Privacy: (1) the content of their resume. (2) the content of their resume and their digital footprint (personal online information)]***
>
> ***[Accountability: (1) Only resumes that meet a certain score are sent to a hiring manager for further review. (2) Together with the score, the resumes are sent to a hiring manager for further review.]***

The complete set of vignettes consists of the 8 full-factorial treatment conditions, whereas the respondents will be randomly assigned to one of the conditions. The assignment to each group will occur with equal probability; therefore, approximately 138 individuals will be assigned to each group, per sample. They will then be tasked to rate the fairness, effectiveness and their general acceptance regarding the specific vignette. The questions to measure the three outcome variables are:

Fairness *[Bipolar 11-point scale: (0) Not fair at all; (10) Very fair]*	*How fair do you think this type of program would be to people applying for jobs?*
Efficiency *[Bipolar 11-point scale: (0) Not effective at all; (10) Very effective]*	*How effective do you think this type of program would be to people applying for jobs?*
Acceptance *[Bipolar 11-point scale: (0) Not acceptable at all; (10) Completely acceptable]*	*Overall, how acceptable do you think it is for public authorities to use this type of program when hiring?*

For the analysis of heterogeneous effects among individuals, we collect a set of covariates (Table 4). Besides socio-demographic factors we measure the political ideology of individuals, their cultural worldviews, as well as the individual level of technophilia. To analyse the relationship between experienced discrimination and acceptance, we also ask respondents on their personal experiences with discrimination, as well as their perceived reasons for this kind of discrimination. For this, we will employ the everyday discrimination short-scale developed by Sterntahl et.al (2011).

Analysis

Since both are forms of factorial survey experiments, the vignette experiment will be analysed in a similar way as the conjoint experiment by estimating the marginal means (MM) for each factor level. This allows estimating the marginal differences in the outcome variables, depending on the presence of the different factor levels. Following the recommendations from Hainmueller et al. (2014), we estimate the MM by using linear probability models and regress the dependent variable on dummy variables for each factorial level. Besides the baseline effects for each factor level, we will also estimate the interaction effect between different factors. This allows us to determine the conditions under which individual acceptance, as well as perceived efficiency and fairness of the hiring process are the highest. To estimate these effects we use the "cregg" package in R with clustered standard errors by respondents, to account for correlations within responses from a given respondent. For the subgroup analysis for heterogeneous effects, we will transform the covariates into categorical factors and group the baseline results from the conjoint analysis by the different covariate groups (e.g. technophile/technophobe, personal experience with discrimination). While the applied quotas of our design aim to account for differences in inclusion probabilities, sampling errors and possible non-responses might still occur, therefore we will use the census data to create post-stratification weights in order to adjust the sample. We will report both the weighted and unweighted results. For testing the robustness of our results, we will conduct f and t-tests.

Appendix

Table 1: Demographic distributions in statistical populations and selected quotas for representative sample Sources: national census statistics (DESTATIS, U.S. Census Bureau, Statistics Bureau of Japan)

Demographic variables	Statistical distribution and quotas per country. Statistical population behind country name (above 18 years old in possession of citizenship)					
	Germany (n=1.100) (61.688.485)		**Japan** (n=1.100) (105.302.800)		**United States** (n=1.100) (255.200.373)	
male	48,33 %	532	48,21 %	530	48,73 %	536
female	51,67 %	568	51,79 %	570	51,27 %	564
age 18-29 (US 18-24)	14,84 %	163	13,39 %	147	11,84 %	130
age 30– 44 (US 25 – 44)	20,30 %	223	21,22 %	233	34,33 %	378
age 45 – 59 (US 45-64)	28,57 %	314	24,38 %	268	32,65 %	359
age over 60 (US over 64)	36,29 %	399	41,01 %	451	21,18 %	233
Education I	33,74 % (Hauptschule and below)	368	14,77 % (less than high school graduate)	163	39,00 % (high school degree or less)	429
Education II	24,37 % (Realschule)	268	39,67 % (high school graduate)	436	28,00 % (some college/ associate degree)	308
Education III	31,93 % (Gymnasium and higher)	351	45,56 % (more than high school graduate)	501	33,00 % (tertiary degree)	363
State or Region						
SH HH NI HB NW HE RP BW BY SL BE BB MV SN ST TH	3,67% 2,10% 9,93% 0,77% 21,36% 7,15% 4,99% 12,53% 15,44% 1,26% 4,06% 3,33% 2,15% 5,40% 3,01% 2,86%	40 23 109 8 235 79 55 138 170 14 45 37 24 59 33 32				
Hokkaido Tohoko Kanto Chubu Kansai Chugoku Shikoku Kyushu/Okinawa			4,29% 7,03% 34,40% 16,66% 17,66% 5,77% 2,99% 11,20%	47 77 378 183 194 63 33 123		
Northeast Midwest South West					17,43% 20,76% 38,05% 23,76%	192 228 419 261

Table 2: Attributes and levels of the conjoint experiment

Attributes	Levels
Safety approval by	public authority manufacturers / industry self-regulation independent body of scientific experts
Safety standards for self-driving cars compared to conventional cars are	stricter much more strict identical
Legal liability for accidents caused by the autopilot	manufacturer car owner
Autopilots ethical prioritization	protect occupant protect others driver decides
Access of telemetric data	only driver driver and public authorities (e.g. traffic control) driver and private companies (e.g. manufacturer, insurance)
Supervision of the autopilot by driver	No supervision Constant supervision Supervision after warning signal by autopilot

Figure 1: Decision-screen for respondents

(Task 1/6)

Please carefully review the two legislative proposals.

	Proposal A	Proposal B
Autopilots ethical prioritization	driver decides	driver decides
Legal liability for accidents caused by the autopilot	manufacturer	manufacturer
Access of telemetry data	driver and private companies (e.g. manufacturer, insurance)	only driver
Safety standards for self-driving cars compared to conventional cars are	much more strict	stricter
Safety approval by	independent body of scientific experts	independent body of scientific experts
Supervision of Autopilot by driver	supervision after warning signal by autopilot	supervision after warning signal by autopilot

Explanation Help

Which proposal for the regulation of self-driving cars do you prefer?

Proposal A Proposal B

Table 3: Conjoint introduction to respondents

A growing number of self-driving cars already participate in regular traffic. Self-driving cars do not require the input of a human driver when they are driving in autopilot mode. In the near future, self-driving cars and cars driven by humans will interact on public roads. This raises the question how self-driving cars should be regulated politically.

We would like to know from you, which regulation for self-driving cars you prefer. For that reason, we will show you tables comparing different legislative proposals. Each proposal consists of a set of rules towards the regulation of self-driving cars. The tables contain the same areas of regulation, however the specific rules for each area change between proposals. You can find an explanation of the areas in the table below:

Area of regulation	Description
Safety approval by	The safety approval is the permission to operate a specific type of car on public roads. Who should be responsible to issue the safety approval for self-driving cars?
Safety standards for self-driving cars compared to conventional cars are	For getting the safety approval, the specific type of car needs to comply with safety standards, for example the probability that fatal accidents occur. How should the safety standards for self-driving cars be set in comparison with conventional cars?
Legal liability for accidents caused by the autopilot	Consider a situation in which a self-driving car in autopilot mode caused a traffic accident. Who should generally bear the costs resulting from the damage, unless another party is proven to be at fault?
Autopilots' ethical prioritization	How should self-driving cars behave in unavoidable collisions with personal injury? Should the autopilot prioritize the protection of the occupant at any cost, should it protect other traffic participants, even at the expanse of the occupants, or should the driver be able to determine who should be protected ?
Access to telemetric data	Telemetric data provides information on the location and speed of vehicles. Self-driving cars collect comprehensive telemetric data. Who should have access to this data?
Supervision of the autopilot by driver	Which role should the driver have, if the autopilot is activated? Should the driver be obligated to continuously monitor the actions of the autopilot, should such obligation only be mandatory after a warning signal by the autopilot, or should the driver not be obligated to monitor the autopilot at all?

In total, six tables comparing two different proposals each, are displayed. Each time you decide which proposal you prefer. You will find the introduced areas of regulation on the very left side, and two different proposals with different rules for those areas in the columns on the right. We illustrate this in the following example . The order of the table might look different for you, than it does in the given example:

Please carefully review the set of features (or rules) of both proposals for all areas of regulation, and select the proposal you overall prefer, by clicking on "Proposal A" or "Proposal B". If you prefer some features of Proposal A, and some features of Proposal B, please select the one with the features that are more important to you.

Table 4: Measured covariates

Theoretical construct	Question Text	Outcome Measure	Source/reference
Political ideology (economic dimension)	What position do you take on taxes and social security? Please use this scale [Lower taxes, although this results in less social services.; More social services, although this results in raising taxes.]	Bipolar 11-point scale	(Roßteutscher et.al 2017)
Political ideology (cultural dimension)	What position do you take on immigration of foreigners? Please use this scale [Immigration for foreigners should be easier. ; Immigration for foreigners should be more difficult.]	Bipolar 11-point scale	(Roßteutscher et.al 2017)
Technophile attitude syndrome	Now we would like to know your opinion on the role of technology in our society. To what extent do you agree or disagree with the following statements? 1. No one can stop the technological progress. 2. In order to preserve the environment, it is necessary for us to constrain our consumerism. (-) 3. The further technology advances, the more compulsive it becomes for people.(-) 4. The technological progress will mean a higher standard of living for future generations. 5. Technology creates more problems than it solves. (-) 6. Technological progress must not be set any limits. 7. The technological developments will help us to solve humanities biggest problems, like hunger, poverty or climate change. 8. It is fashionable to use the latest technological devices. 9. All problems can be solved by the use of technology.	Battery of 9 items; 5 point-scale [Strongly disagree, Somewhat disagree, Neither agree nor disagree, Somewhat agree, Strongly agree]; statements with a (-) will be reversed, order randomized	(Technik Radar 2018)
Cultural cognition worldview scales (shortscale) A) Group or Individualism-Communitarianism	People in our society often disagree about how far to let individuals go in making decisions for themselves. How strongly you agree or disagree with each of these statements? 1. The government interferes far too much in our everyday lives. 2. Sometimes the government needs to make laws that keep people from hurting themselves. (-) 3. It's not the government's business to try to protect people from themselves. 4. The government should stop telling people how to live their lives. 5. The government should do more to advance society's goals, even if that means limiting the freedom and choices of individuals.(-) 6. Government should put limits on the choices individuals can make so they don't get in the way of what's good for society.(-)	Battery of 6 items; 6 point-scale [Strongly disagree, Moderately disagree, Slightly disagree,Slightly agree,Moderately agree, Strongly agree]; statements with a (-) will be reversed, order randomized	(Kahan 2012)
Cultural cognition worldview scales (shortscale) B) Grid or Hierarchy - Egalitarianism	People in our society often disagree about issues of equality and discrimination. How strongly do you agree or disagree with each of these statements? 1. We have gone too far in pushing equal rights in this country. 2. Our society would be better off if the distribution of wealth was more equal.(-) 3. We need to dramatically reduce inequalities between the rich and the poor, whites and people of colour, and men and women.(-) 4. Discrimination against minorities is still a very serious problem in our society.(-) 5. It seems like blacks, women, homosexuals and other groups don't want equal rights, they want special rights just for them. 6. Society as a whole has become too soft and feminine.	Battery of 6 items; 6 point-scale [Strongly disagree, Moderately disagree, Slightly disagree,Slightly agree,Moderately agree, Strongly agree]; statements with a (-) will be reversed, order randomized	(Kahan 2012)

Everyday discrimination short-scale (Measures)	In your day-to-day life how often have any of the following things happened to you? 1. You are treated with less courtesy or respect than other people. 2. You receive poorer service than other people at restaurants or stores. 3. People act as if they think you are not smart. 4. People act as if they are afraid of you. 5. You are threatened or harassed 6. You have been unfairly fired from a job, unfairly denied a promotion or unfairly not been hired for a job	6 response categories: • Almost everyday • At least once a week • A few times a month • A few times a year • Less than once a year • Never	(Williams et.al 1997, Sternthal et.al 2011)
Everyday discrimination short-scale (Follow-up) (Asked only of those answering "A few times a year" or more frequently to at least one question.)	What do you think is the main reason for these experiences?	11 response options • Your Ancestry or National Origins • Your Gender • Your Race • Your Age • Your Religion • Your Height • Your Weight • Some other Aspect of Your Physical Appearance • Your Sexual Orientation • Your Education or Income Level • A physical or mental disability	(Williams et.al 1997, Sternthal et.al 2011)

Literature

Auspurg, K and Hinz, T. (2014).Factorial Survey Experiments. ThousandOaks, CA: Sage

Bechtel, M; Hainmueller, J & Margalit, Y (2012): Studying Public Opinion on Multidimensional Policies: The Case of the Eurozone Bailouts. , 2012,

Bonnefon, J.-F., Shariff, A. & Rahwan, I. (2016) The social dilemma of autonomous vehicles. Science 352 (6293), 1573–1576.

Burton, JW, Stein, M-K, Jensen, TB.(2020). A systematic review of algorithm aversion in augmented decision making. J Behav Dec Making. 2020: 33: 220– 239.

Choi, J. K. & Ji, Y. G. (2015) Investigating the Importance of Trust on Adopting an Autonomous Vehicle. *International Journal of Human–Computer Interaction* 31 (10), 692–702.

Chong, D & Druckman, J (2007). Framing Theory. Annual Review of Political Science, 10 , 103–126.

Chong D and Druckman JN (2012) Counterframing effects. The Journal of Politics 75(1): 1–16.

Geels, F. W. & Schot, J. (2007) Typology of sociotechnical transition pathways. *Research Policy* 36 (3), 399–417.

Gkartzonikas, C. & Gkritza, K. (2019) What have we learned? A review of stated preference and choice studies on autonomous vehicles. Transportation Research Part C: Emerging Technologies 98, 323–337.

Hainmueller, J, Hopkins D.,Yamamoto, T (2014). Causal Inference in ConjointAnalysis."PoliticalAnalysis 22(1):1–30

Hainmueller, J, Hangartner D, Yamamoto, T (2015) "Validating Vignette and Conjoint Survey Experiments against Real-world behavior."ProceedingsoftheNationalAcademyofSciences 112(8):2395–2400.

Kahan, D. M. (2012). Cultural cognition as a conception of the cultural theory of risk. In S. Roeser (Ed.), Handbook of risk theory (pp. 725-759). Springer Netherlands. Available at SSRN: http://ssrn.com/abstract=1123807.

Majone, G. (1975) On The Notion Of Political Feasibility. *European Journal of Political Research* 3 (3), 259–274.

Montpetit, É. (2008) Policy Design For Legitimacy: Expert Knowledge, Citizens, Time And Inclusion In The United Kingdom's Biotechnology Sector. *Public Administration* 86 (1), 259–277.

Pavlou, Paul A. (2003) Consumer Acceptance of Electronic Commerce: Integrating Trust and Risk with the Technology Acceptance Model. International Journal of Electronic Commerce 7 (3), 101–134.

Reins, L (2019). Regulating New Technologies in Uncertain Times – Challenges and Opportunities. In Reins, L (Ed.) Regulating New Technologies in Uncertain Times. 3-18. Information Technology and Law Series, Vol. 32. T.M.C. Asser Press, The Hague.

Rip, A., Kemp & René (1998) Technological change. *Human choice and climate change. Vol. II, Resources and Technology*, 327–399.

Roßteutscher, Sigrid; Schoen, Harald, Schmitt-Beck, Rüdiger; Weßels, Bernhard; Wolf, Christof; Wagner, Aiko (2017): Nachwahl-Querschnitt (GLES 2017). GESIS Datenarchiv, Köln: ZA6801 Datenfile Version 2.0.0, doi: 10.4232/1.1299

Savulescu, J., Kahane, G. & Gyngell, C. (2019) From public preferences to ethical policy. *Nature Human Behaviour* 3 (12), 1241–1243.

Shariff, A., Bonnefon, J.-F. & Rahwan, I. (2017) Psychological roadblocks to the adoption of self-driving vehicles. Nature Human Behaviour 1 (10), 694–696.

Smith, A., Stirling, A. & Berkhout, F. (2005) The governance of sustainable socio-technical transitions. *Research Policy* 34 (10), 1491–1510.

Smith, A. (2018) Public Attitudes Toward Computer Algorithms (Pew Research Center, 2018).

Sniderman PM, Theriault SM. (2004). The structure of political argument and the logic of issue framing. In Studies in Public Opinion, ed. WE Saris, PM Sniderman, pp. 133–65. Princeton, NJ: Princeton Univ. Press

Sternthal, M., Slopen, N., Williams, D.R. (2011) "Racial Disparities in Health: How Much Does Stress Really Matter?" Du Bois Review, 2011; 8(1): 95-113.

TechnikRadar 2018 (Version: 1.0.0). , DOI: https://doi.org/10.7802/1904,

Wallner, J. (2008) Legitimacy and Public Policy: Seeing Beyond Effectiveness, Efficiency, and Performance. *Policy Studies Journal* 36 (3), 421–443.

Wiener JB. (2004) The regulation of technology, and the technology of regulation. Technology in Society 2004;(26):483-500.

Williams, D.R., Yu, Y., Jackson, J.S., and Anderson, N.B. (1997) "Racial Differences in Physical and Mental Health: Socioeconomic Status, Stress, and Discrimination." Journal of Health Psychology. 1997; 2(3):335-351.