Chapter 3: Data and Methods

A. Data

This work draws on data from the project *Justice, an individual, contextual or situational affair?* (JInCS) which were specifically collected for the purpose of enabling a disentanglement of the factors contributing to our perceptions of distributive justice. The project aims to contribute towards a deeper understanding of perceptions of distributive justice. JInCS was fully funded by the Doc.CH program of the Swiss National Science Foundation (SNSF — P0BEP1_168648). Doc.CH is a support program, financing young academics in the Humanities and Social Sciences for the duration of their PhD studies.

The data stemming from the JInCS project were collected from three different samples. While it was originally planned as a general population survey in Switzerland, including all language-regions, it was then expanded to include two student samples. The student data were collected at the University of Bern, Switzerland and Princeton University in the United States. Prior to data collection, both quantitative and qualitative pretests were performed in October and November 2017.

I. Samples

While the data for the Swiss general population were collected using both paper and pencil as well as online questionnaires — the respondents chose what suited them best — the student data were solely collected using an online questionnaire. While it was cheaper and more efficient to collect data digitally, the respondents from the general population sample had to be contacted via letters sent through the mail, because home addresses were the most reliable source of contact data. Because the Swiss general population data included all age groups, contacting respondents by email was not a realistic option.

1. General Population Sample in Switzerland

The respondents from the Swiss general population sample were drawn from a stratified random sample (using region, gender and age as stratifying variables) by MS Direct, a marketing company with a large address register. The Italian-speaking region was oversampled to allow for a differentiation of context-level effects within Switzerland[1]. In late 2017, when the contact data of the respondents were drawn, MS Direct had over 3 million entries. In a population of approximately 8.4 million at that time, this corresponds to over a third of the Swiss population. The data were collected between February and June 2018.

To raise response rates, Dillman's (2007) tailored design method (TDM) was closely followed. TDM is based on social exchange theory and the assumption that by reducing the costs for the participants and increasing the benefits of participation, they will be more likely to cooperate and provide researchers with their data (Dillman 2007). To establish trust and try to convince participants of the relevance of their contribution, participants were contacted three times. First, they received a letter announcing that they had been selected at random to participate in the survey. The second letter contained six items. Next to the paper-pencil (PAPI) questionnaire, these were: a letter to inform the participants on how to proceed, including an individual code for online participation; a post-paid envelope to reduce the costs of participating; a card they could send back with their contact data if they wanted to be informed of the results; as an incentive, they received two customised printed webstamps worth CHF 1.00 each; as well as a thank you note to which the stamps were attached. About two weeks later, the respondents were contacted once again to thank those who had already participated and to remind those who had not yet gathered the willingness or had time to respond of the importance of their participation. On the whole, with a response rate of 40%, the measures seem to have paid off. Of the approximately 4500 people contacted, 1.789 people responded.

1 While there are slight differences, these within-country context effects will not be discussed further here.

2. University of Bern

The Swiss student data were collected at the University of Bern in February and March 2018. The students were invited to participate via email. Because of strict data protection regulations, the emails were sent out by the admissions office (*Zulassung, Immatrikulation und Beratung*). Also due to the strict rules, it was not permitted to contact the students more than once. Of the 12.735 (doctoral) students who received the email (everyone enrolled that term) 1.081 students participated. A reason for the rather low response rate of 8.5% might have been that Unipark, the tool used for the survey, allows for only 100 people to be active in the survey simultaneously and the time point of the highest response rates was more or less immediately upon receipt of the email.

3. Princeton University

The data collection at Princeton University was kindly organised by Edward Freeland, associate director at the Survey Research Center (SRC), Princeton University. Email invitations were sent out to all 2.910 enrolled graduate students at Princeton. Of these, 358 students participated in the online questionnaire, yielding a response rate of roughly 12%. The data were collected between March and April 2018.

In total, 1.439 students from the University of Bern and Princeton University participated in the survey. Together with the respondents from the Swiss general population that makes up a total of 3.228 respondents or, since each of distributional survey experiments consisted of three outcomes, 9.684 observations for the experiments.

B. Methods

In this section, first survey experiments in general are discussed, before the distributional survey experiment (DSE) is introduced as a new approach. In a next step, the designs of the four DSEs are presented before delving into the question of analytical strategies.

I. Survey Experiments

Since questions of distributive justice can be considered sensitive issues from an ethical and political point of view, social desirability bias can be a problem worth thinking about when collecting the data. Survey experiments can be a very powerful tool for the reduction of social desirability bias (Auspurg, Hinz, Liebig, et al. 2015). First, through the hypothetical nature of the question, respondents feel less threatened (Badgett and Folbre 2003). Second, the dimensions/variables of interest are not brought to the attention of the respondents explicitly, but indirectly as part of a whole description. This makes respondents less aware of the experimental context. Evaluating characteristics of interest in bundles, as is the case with survey experiments, makes the situation closer to life, while also lowering the potential for social desirability bias (Alexander and H. J. Becker 1978; Badgett and Folbre 2003; Wallander 2009). Furthermore, another advantage of survey experiments is that they are a highly efficient data collection method. The researcher can manipulate several treatments in one and the same experimental setup. So as not to overburden the respondent, researchers should restrict the number of variables to seven (+/-2) (Auspurg and Hinz 2015). Additionally, survey experiments combine the advantages of conventional population-based surveys and experiments (Angrist and Pischke 2009; R. Becker and Zangger 2015). Since survey experiments are integrated in a questionnaire, they are more economical than experiments in laboratory settings and can reach more people, thus raising external validity. Also, because they use efficient designs, in which the dimensions that make up a vignette are brought into a nearly orthogonal array, survey experiments have the valuable property of circumventing the problem of multicollinearity of different dimensions (Rossi and A. B. Anderson 1982; Badgett and Folbre 2003). Because of these design properties, it is possible to compute the relative importance of the single dimensions for the overall assessments (Rossi and A. B. Anderson 1982).

II. The Distributional Survey Experiment

The distributional survey experiment incorporates all the above mentioned useful properties but the design is more suited to capturing the essence of the problem inherent in allocation decisions. The design of the DSE combines some of the useful features of a factorial survey experiment, a choice experiment and game experiments in laboratories. While in a factorial

survey experiment (FSE) respondents are presented one vignette at a time and asked to make ratings on a scale (Auspurg and Hinz 2015), in choice experiments (CE) respondents are shown a collection of typically three vignettes as a group, usually presented in the form of a table, and they are asked to select the alternative they think is best (Street, Burgess, and Louviere 2005; Kuhfeld 2005).

The DSE combines the advantageous properties of both approaches but adds a new element that makes it better suited for the issue at hand: By asking the respondents to distribute a prespecified sum among three[2] people described in vignettes. This is a more active task that involves the weighing up of justice principles across vignettes and thus captures the relevant evaluative processes people undergo when attempting to make a just distribution better than in a FSE or a CE. The DSE brings the advantage of incorporating the fact that in real life, given that resources are limited and sometimes scarce, people making allocation decisions are faced with the necessity of making trade-offs between competing claims. It is rarely possible to give to everyone in such a way so as to, for example, simultaneously satisfy both need an merit claims of different people; when resources are limited, giving more to one person, automatically leaves less for the other two. This means that people need to prioritise, and that what the three people receive is inherently interdependent.

To account for this interdependence, the DSE is designed in the same way as a choice experiment, meaning that the vignettes were constructed using a D-efficient arrangement and that the single vignettes were additionally grouped efficiently into choice sets (Kuhfeld 2005). Therefore, the dimensions/variables making up the vignettes are arranged in a way that maximises orthogonality and optimises level-balance. Visually, and perhaps task-wise, because it forces respondents to integrate the information from three vignettes in the making of one decision, the DSE is more similar to a choice experiment than a factorial survey experiment. However, in a choice experiment we would only be left with the information on which of the three alternatives was most attractive for the respondent — in our case, this would correspond to the answer given in response to a question like: "Who would you want to give the reward to?" — and we would not know which of the other alternatives ranked second or third and by how much. The DSE leaves us with more abundant information on the preferences of the respondents.

2 Of course this number can be modified, however, in choice experiments it is most common to group vignettes in choice sets of three.

While the outcomes are metric, just as is usually the case with FSEs, the difference is that instead of independent ratings for each "vignette person" — in our case, this would correspond to the answer to a question like: "How deserving is this person of a reward?"[3] — the outcome is the sum each person described in a vignette receives, given the combination of attributes of the other two people. This is the biggest advantage of the DSE compared to other methods. Because of the inherent interdependence of the three outcomes per DSE, we are doing more justice to the nature of the problem than would be possible with either a factorial survey experiment or a choice experiment.

As in real life allocation problems, respondents are faced with a complex situation, in which they need to weigh up the attributes of one person against those of another and decide who should get how much based on their sense of justice. This is complex, because even within one person we might have conflicting information, so that in some regards the person can be seen as deserving on the grounds of merit and in others less so, and the same goes for variables depicting the person's need. And then, one person who is needy might not also be very deserving. Furthermore, these evaluations need to be performed keeping the relevant criteria of the competitors (the other two people described in the DSE) in mind. Additionally, factors that have nothing to do with merit or need, such as ethnic origin or gender, could have an effect on justice evaluations, even if subconsciously.

Summing up, the DSE creates a situation in which respondents evaluate a set of people (here three) — each of whom has their own individual combination of attributes — simultaneously and in direct comparison with the others in the set. As in a real life allocation problem, one person's gain is another's loss. In real life, problems of distributive justice become especially salient in cases where we feel that scarce resources are not being justly distributed. What do we do when we have less resources than what would be necessary to make ideal distributions according to our individual sense of justice? We are forced to prioritise and an efficient way to do this is to take interactions of need and merit variables into consideration — this is possible with a DSE. Additionally, as in a real distribution problem, respondents have to make an implicit relative positioning of the people described in the vignettes in comparison with the other two people in the choice set. Compared to a FSE or a CE, a DSE is arguably psychologically more satisfactory for the

3 Or, as has been done frequently (Auspurg, Hinz, and Sauer 2017), respondents are asked to rate the fairness of a described person's wage, for example, by rating it on scale from *much too high* to *much too low*.

respondents, since it allows them to make more detailed differentiations in line with their preference order. For example, they can distribute something (or even the same amount) to everyone, if that is what they think is the fairest option. It is not possible to capture preferences this precisely using FSEs or CEs, especially because these methods do not allow for interdependencies to be accounted for to this degree. The case for the DSE becomes especially apparent when we think of radical egalitarians, for whom it is never possible to reveal their preference when forced to make a justice evaluation on a scale or to pick one individual over the other two. The DSE offers respondents much more freedom to express their underlying preferences. For these reasons, the DSE is a powerful tool for the collection of data on distributive justice. It provides the researcher with data that depict the real-life trade-offs that people are forced to make when distributing goods according to their justice preferences more precisely compared to alternative methods. On top of that, the flexibility of the DSE enables us to measure and take into consideration the complex interplay of individual-level effects, the effects of context and the situation.

III. Designs of The Distributional Survey Experiments in Four Situations

To include situation as a relevant factor in justice evaluations, four DSEs were designed in different settings. Money was to be distributed among friends, at the workplace, among family members and among prospective students applying for a scholarship. While in general any kind of resource that can be subdivided is a possibility, monetary resources are convenient. For one thing respondents are familiar with using money in allocation problems and for another, money allows them much freedom in subdividing the amount into shares as proportionally small or large as they please. The sum to be distributed varied randomly between the choice sets, with either CHF 3000.00, CHF 9000.00, CHF 18000.00 or CHF 30000.00 [4]. In all DSEs, all two-way interactions between dimensions were included in the design. In the following, each of the distributional survey experiments will be presented separately. They will be presented in the order as they appeared in the questionnaire. Additionally, a short introductory text explaining the task, shown in figure 3.1, was placed immediately preceding the first DSE:

4 Different sums were used so as to allow for an evaluation of the effects of the size of the sum. However, this will not be our focus here.

In the following 4 questions, you can distribute a specified amount of money among the people described in the table. You can give one person everything and nothing to the other two. Of course, you could also give everybody the same share or distribute the money any other way. You decide!

In any case, everyone will be asked to keep the received amount a secret, so no one gets jealous.

Figure 3.1: Introductory text to the DSEs

1. The DSE in the Friends Situation

In the DSE representing the friends setting, the vignettes were constructed using nine dimensions with two or three levels each[5]. In the end, a setup consisting of 144 vignettes arranged in 48 choice sets of groups of three was found. Table 3.1 provides an overview of the variables included in the DSE and their levels.

Characteristics	**Values**
Gender	1 male, 2 female
Ethnic origin	1 local name, 2 Slavic name, 3 Arabic name
Need	
Private life	1 lives with partner, 2 single
Children	1 no children, 2 two children
Health	1 in good health, 2 in poor health
Financial situation	1 easy, 2 tight
Merit	
Reliability	1 a good friend when times are good, 2 you can count on them when times are rough
Trustworthiness	1 you better watch out what you tell them, 2 you can trust them with anything
Friends since	2 years , 8 years , 14 years

Table 3.1: Characteristics included in the DSE in the friends situation

Four of the variables operationalise need, while three of the variables are merit indicators. Additionally, the information on gender (male or female) and ethnic origin is provided through the name, which was chosen to sound either local (depending on the context either Swiss or American), Slavic or

5 In the literature on survey experiments, two levels are recommended, unless more is necessary due to the particular research interest or if non-linear effects are expected (Auspurg and Hinz 2015).

Arabic. The names chosen to represent local (Swiss or American) names are: "Sarah" and "Ben". "Leila" and "Youssef" are used as the male and female Arabic names while "Jelena" and "Nenad" represent a Slavic ethnic background. Need is operationalised using information on whether a person lives with their partner or whether they are single; whether they have children or not and whether they are in good or poor health. The assumptions were that people living with a partner, those without children and those who enjoy good health *need* less, because they can share expenses, do not have dependants and do not have special health related costs. In the combination: single with children, the person was described as a "single parent" instead of just "single". Additionally, a "tight financial situation" indicates a higher need than an "easy financial situation". Merit is operationalised through reliability and trustworthiness as well as by how long they have been friends for. A friend was described as reliable if "you can count on them when times are rough" as opposed to only "when times are good". A trustworthy friend was described as someone "you can trust with anything", while if "you better watch out what you tell them", the friend is not very trustworthy. The length of the friendship is set to 2, 8 or 14 years.

In some of the analyses, the merit variables were combined in one variable constructed as an additive scale with values ranging from 3 to 7, depending on the levels of the single merit indicators reliability, trustworthiness and years of friendship. The story introducing this DSE was that respondents were asked to imagine having won a large sum in the lottery and that to celebrate, they were going to share a portion of their win among three of their friends. An example DSE in the friends setting is shown in figure 3.2.

2. Imagine you actually won quite a bit of money playing lotto and – to celebrate – you would like to buy your friends Nenad, Sarah, and Youssef presents for a total of CHF 18'000. Who would you spend how much on?

	Nenad	Sarah	Youssef
private life	lives with partner	lives with partner	single
children	no children	2 children	no children
health	in poor health	in good health	in good health
financial situation	tight	easy	easy
as a friend	you can count on him when times are rough you better watch out, what you tell him	you can count on her when times are rough you can trust her with anything	a good friend when times are good you better watch out, what you tell him
friends since	14 years	2 years	8 years
sum	+	+	= 18'000

Figure 3.2: DSE in the friends situation

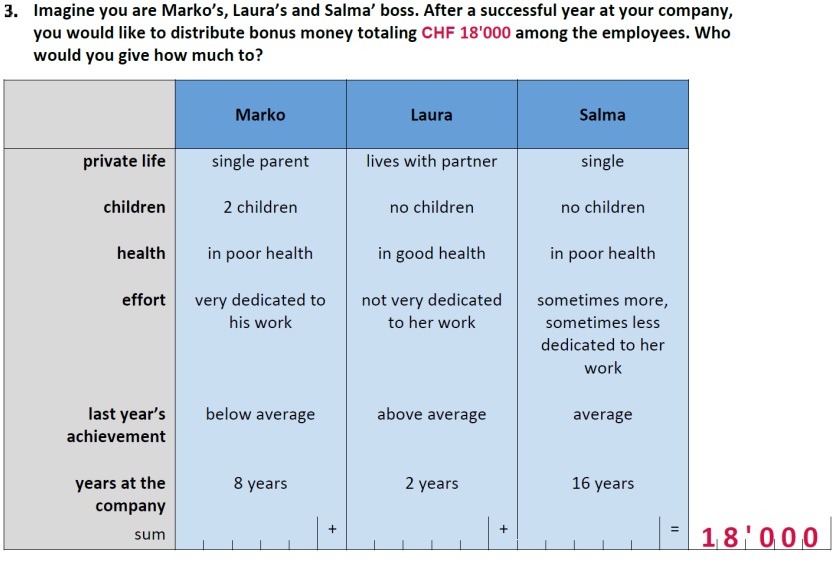
2. The DSE in the Work Situation

The second distributional survey experiment was constructed using eight dimensions with two to three levels each. From the vignette universe, a subset of 216 vignettes was selected and arranged into 72 different choice sets. The dimensions and their levels are summarised in table 3.2.

As in the above described DSE, gender (male, female) and ethnic origin (local, Slavic, Arabic) are operationalised through the names: "Laura" and "Lukas/Luca" for the local version, "Salma" and "Ahmad" (Arabic) and "Milena" and "Marko" (Slavic). Furthermore, information on the merit and need of the hypothetical employees was provided. Need is operationalised using three of the same variables used in the first DSE: whether a person lives with their partner or is single, whether they have children or not, and the state of their health.

Characteristics	Values
Gender	1 male, 2 female
Ethnic origin	1 local name, 2 Slavic name, 3 Arabic name
Need	
Private life	1 lives with partner, 2 single
Children	1 no children, 2 two children
Health	1 in good health, 2 in poor health
Merit	
Dedication	1 not very dedicated, 2 more or less dedicated 3 very dedicated
Performance	1 under average, 2 average 3 above average
Years at the company	2 years , 8 years , 16 years

Table 3.2: Characteristics included in the DSE in the work situation



3. Imagine you are Marko's, Laura's and Salma' boss. After a successful year at your company, you would like to distribute bonus money totaling CHF 18'000 among the employees. Who would you give how much to?

	Marko	Laura	Salma
private life	single parent	lives with partner	single
children	2 children	no children	no children
health	in poor health	in good health	in poor health
effort	very dedicated to his work	not very dedicated to her work	sometimes more, sometimes less dedicated to her work
last year's achievement	below average	above average	average
years at the company	8 years	2 years	16 years
sum	+	+	= 18'000

Figure 3.3: DSE in the work situation

Merit is operationalised with information on dedication to job ("not very", "more or less" or "very" dedicated to job), performance levels ("under average", "average" or "above average") as well as duration of employment (2, 8 or 16 years). For some of the analyses, the variables operationalising merit were combined in one additive scale variable with values ranging from

3 to 9, depending on the levels of the single variables. In the work situation, respondents were asked to imagine they were the boss of the employees described in the DSE. The money involved was said to be an end of year bonus. Figure 3.3 shows an example DSE for the work setting.

3. The DSE in the Family Situation

The third DSE is set in a family context. More specifically, participants were asked to divide an inheritance among their hypothetical grown-up children. The vignettes in the family DSE consist of six dimensions with two levels each except for one three-level variable.

Characteristics	Values
Gender	1 male, 2 female
Need	
Private life	1 lives with partner, 2 single
Children	1 no children, 2 two children
Health	1 in good health, 2 in poor health
Financial situation	1 easy, 2 tight
Merit	
Attentive & helpful	1 not very, 2 mostly, 3 very

Table 3.3: Characteristics included in the DSE in the family situation

The design is summarised in table 3.3. As in the other DSEs, information on the gender of the hypothetical family member was provided through the name. However, since the "vignette people" in this DSE are all members of the same family, the information on ethnic origin was not added as an additional dimension but kept constant at a local name. "Anna/Anne" or "Sylvia" were used for women and "Lorenz/Laurence" or "David" for men. Need was operationalised using the same four variables as in the first DSE indicating: whether they live with a partner or are single; whether they have children or not; their health status as well as their financial situation (tight or easy). In this DSE merit was operationalised using only one variable: the degree of attentiveness and helpfulness in the past. This variable has three levels describing whether the family member was very, mostly, or not very attentive and helpful in the past. This DSE consists of a subset of 96 vignettes, which were grouped into 32 choice sets. Figure 3.4 shows an example DSE in the family situation.

4. Imagine having to distribute CHF 3'000 inheritance among the siblings Anne, Sylvia and David. How would you distribute the money if they were your grown-up children?

	Anne		Sylvia		David		
private life	single		lives with partner		single		
children	no children		2 children		no children		
health	in poor health		in good health		in poor health		
personality	highly considerate and has always helped out in the past when needed		not very considerate and has not been very helpful in the past		usually considerate and has often helped out in the past		
financial situation	easy		tight		easy		
sum		+		+		=	3'000

Figure 3.4: DSE in the family situation

4. The DSE in the Public Goods Situation: Scholarships

In the fourth DSE portraying a public goods situation, respondents were asked to divide scholarship money among three hypothetical students. Table 3.4 provides an overview of the dimensions and levels used to describe them. In total, eight variables, with either two or three levels each, provided information on the gender, ethnic background as well as the need and merit of the "vignette people".

The names used to operationalise gender and ethnic origin were: "Daniela/Danielle" and "Stefan/Steven" (local), "Amal" and "Amir" (Arabic) and "Lidija" and "Matia" (Slavic). Need was operationalised using parental education, set to either compulsory schooling, vocational education and training (VET), or university education, as an indicator of social class. Furthermore, information on the financial situation (described as easy or tight) of the student's parents was also provided. Merit was operationalised through school leaving grades as well as through information on how much effort the student puts into their work. They were described as either "not very hard-working", "more or less", or "very hard-working". Additionally, two dimensions were neither clearly merit nor need indicators. One of these variables provides information on whether the prospective student needs to move for their studies or whether they can continue living with their parents. The other dimension

Characteristics	Values
Gender	1 male, 2 female
Ethnic origin	1 local name, 2 Slavic name, 3 Arabic name
Need	
Parental education	1 compulsory schooling, 2 vocational education and training (VET), 3 university
Financial situation of parents	1 easy, 2 tight
Merit	
Effort	1 not very hard working, 2 more or less hard working 3 very hard working
Grades	4.5, 5.0, 5.5 (Switzerland) or: 2.0, 3.0, 3.7 (United States)
Additional information	
Working student	1 up to 1/2 day per week, 2 1/2 to 1 day per week, 3 1 to 1 1/2 days per week
Relocation	1 can live at home, 2 has to move out

Table 3.4: Characteristics included in the DSE in the public goods (scholarship) situation

that is neither clearly a need nor merit indicator is how many hours the student will spend working at a side job during their studies.

This DSE was constructed using a subset of 216 vignettes that were grouped into 72 choice sets containing three vignettes each. The values of the merit scale variable range between 2 and 5. Figure 3.5 shows an example DSE in the public goods setting. There are two additional features in this distributional survey experiment. For one thing, the mentioned subject was randomly set to either history or medicine. For another, the last sentence in the example in figure 3.5 ("Please keep in mind that, in some cases, your decision greatly influences whether the young adults can afford to study") was only included in about half the cases (also random variation). This sentence was added to test for potential priming effects of the stimulus which was meant to trigger a sense of responsibility in the respondents. However, these two treatments are not directly relevant to the question at hand and will not be used in the analyses[6].

6 Results of preliminary analyses show that including these variables in the analyses only has a minor effect on the results. While the subject seems to have virtually no effect on the distributions, there seems to be a minor effect of the added sentence aiming to make the respondents feel more personally responsible. However, this additional information primarily influences the results indirectly through the side-job variable.

5. Amir, Steven and Lidija have just graduated from high school. They would like to study history and have all applied for a scholarship. How would you distribute the total sum of CHF 9'000 (per year) among the three, if it were for you to decide?

The scholarships are funded by tax money and will be payed to the recipients for the duration of their studies. Please keep in mind that in some cases your decision greatly influences whether the young adults can afford to study.

	Amir	Steven	Lidija	
final grade	5	5.5	4.5	
learning habits	very hard-working	very hard-working	not very hard-working	
for his/her studies	can stay at home	can stay at home	has to move out of parents house	
will work	0-10%	20-30%	10-20%	
parent's highest education	university degree	vocational education and training	compulsory education	
parent's financial situation	tight	easy	tight	
sum	+	+	=	9'000

Figure 3.5: DSE in the public goods / scholarship situation

IV. Analytical Strategies

Since this is the first application of distributional survey experiments, suitable methods for analysis needed to be developed. Because money is distributed among three people described in vignettes, each DSE leads to three metric outcomes. There are at least three strategies we can use to analyse this data and they will each be described briefly in the following.

1. Treating Outcomes as Independent

The most straightforward method for the analysis of a DSE is to treat the three metric outcomes it generates as if they were single factorial survey experiments. Then the amount each of the three people receives in a DSE is

The dimensions operationalising merit and need are not affected in any substantive way.

the dependent variable. In our case, with randomly varying total sums that were distributed, this needs to be accounted for in the analysis. One way to do this is to standardise the values. Before analysis, the data set needs to be reshaped into long format, so that we have three observations per respondent. To account for the fact that one and the same person is responsible for three outcomes per DSE, and that the distributions are thus interdependent, using a multilevel model is advised. This method has the benefit of allowing us to look at the effects expressed in monetary terms.

2. Treating the Outcomes as Ranks

The second approach pays more tribute to the interdependency of the individual outcomes by, for example, ranking the sums from 1 for the highest to 3 for the lowest values. In the case of equal splits, everyone is assigned a shared value, such as e.g. 1. In contrast to the first method, this procedure ensures that the interdependency of the single vignette outcomes in a DSE is respected. Instead of focusing on the individual amounts, in this approach, we focus on the rank positioning of the single vignettes.

3. Treating the Outcomes as Shares

And third, we can analyse the data from the DSE after first converting the sums of the single vignettes within a DSE into shares of the total sum. This method incorporates some useful properties of both the aforementioned analytical strategies. For one thing, information on the rank is included implicitly. For another, this approach enables a higher exploitation of the information compared to using ranks. The advantage of using shares compared to ranks is that the magnitude of the differences between the sums of the single vignettes is taken into account. This method incorporates the available information in a maximally efficient manner, while accounting for the interdependency of the single outcomes. This is ensured in that the shares, just like ranks, include information on the relative position of each vignette outcome compared to the other two vignettes in a choice set making up a DSE. It is thus the method with the highest statistical power and smallest standard errors. However, most importantly, this procedure is most in line with the theoretical considerations that led to the development of the DSE in the first place. Since all outcomes are treated as inherently interdependent, this method does most justice to a

core property of distributive justice: that allocation problems are usually a zero-sum game. When resources are finite, each individual outcome affects all others and one person's gain is another's loss.

4. Considering the Multilevel Structure

Given that next to the information on the level of the experiment (level 1), additional data were collected on the level of the respondent (level 2), all of the above described strategies require multilevel modelling. Because multilevel and longitudinal data structures rely on the same framework, to take up standard notation, I will speak of *timepoints*. However, in our case, these refer to the fact that we have three observations per person and experiment, which is why we can think of *timepoints* as an indicator for a single vignette. Equation (1) is then a linear multilevel model for a set of individuals i, for which we have observations on at least two "timepoints" t, or in our case three vignettes per distributional survey experiment.

$$y_{it} = \mu_t + \beta x_{it} + \gamma z_i + \alpha_i + \varepsilon_{it} \tag{3.1}$$

The dependent variable y_{it} can be considered interval scaled or ratio scaled when using shares. The constant μ_t is allowed to vary by timepoint; the regression coefficients are given by β and γ for the time-varying predictors x_{it} and the time-constant predictors z_i respectively. We also have two error terms; α_i is a constant that varies over individuals but not over time and ε_{it} that varies randomly by individual as well as by timepoint. Depending on whether we are estimating a fixed effects or random effects model however, we rely on different underlying assumptions for α_i. In the fixed effects models, α_i represents all unobserved time-constant variables and as such is allowed to be correlated with the time-specific predictor variables x_{it} and also with the time-constant variables z_i (Allison 2009). This latter property is at once the main strength and, depending on the research question, also "weakness" of the fixed effects model. Since we are controlling for all individual-specific variation, both observed and unobserved, we can make a causal claim. However, if we are interested in the coefficients for z_i, we have to use other models, since the fixed effects model cannot provide estimates for these level 2 variables. The random effects model offers a solution to this "problem", if we look at it as such, by assuming that α_i represents just another source of random variation: a set of random variables that have a prespecified probability distribution.

This implies that the individual error term α_i is uncorrelated with all other independent variables as well as with the time-specific error term ε_{it}. If this assumption holds true, then we yield unbiased estimates and the model is more efficient than the fixed-effects model. Luckily we can easily test whether this is the case, using either the Hausman test[7] or by comparing the within and between effects of a hybrid model (Allison 2009; Firebaugh, Warner, and Massoglia 2013). If the individual time-constant error term proves to be non-random, the estimates from a random effects model would be biased. The hybrid model is very convenient in that it allows us to combine the best traits of fixed and random effects models, so that we can exploit the efficiency of a random effects model and the consistency of a fixed effects model at the same time. Hybrid models are in essence equivalent to random effects models that control for time-constant unobserved heterogeneity (Schunck 2013). By demeaning or centering the time-variant variables, that is, by calculating unit-specific means and then calculating the deviation scores from the means for the single values, we can decompose the effects of time-varying variables into within and between variance. We can then add these within and between effects separately into our random effects model and also include the time-invariant variables (Allison 2009; Schunck 2013).

For the models shown in chapter 4, random-effects models were always used in the case that they yielded consistent estimates; where this was not the case, fixed effects models were computed[8]. When effects of respondent level characteristics were of theoretical interest, hybrid models were estimated (Rabe-Hesketh and Skrondal 2012; Schunck 2013; Allison 2009).

All the analyses are based on the four distributional survey experiments collected from three different populations: the Swiss general population, students at the University of Bern and students from Princeton University. The aim of the DSEs was to make it possible to differentiate between the principles of merit, equality and need. Additionally, the DSE enables us to collect data efficiently on the individual level (respondent characteristics) and the level of the situation (DSEs in different settings) as well as to include context, by administering it to different populations. Since the distributive survey experiments made up the core element of the JInCS project, they

7 The Hausman specification test can be used to compare the fixed effects estimator, which always yields consistent coefficients, with the random effects estimator, which yields consistent estimates only if there are no significant differences between the two estimators (Hausman 1978; Rabe-Hesketh and Skrondal 2012).

8 This was the case when the Hausman test was positive, thus indicating systematic differences between the fixed effects and random effects models.

were placed at the beginning of the survey, directly after the ice breaker question[9]. Placing the DSEs at the beginning of the questionnaire ensured that, at least in the online version, the most relevant data were secured even if people decided to break off the survey. Additionally, the DSEs were placed early on because they are somewhat more cognitively taxing than the more traditional questions that followed. These were a set of questions on a wide range of issues of distributive justice. Furthermore, at the very end of the questionnaire, respondents were asked a set of socio-demographic questions.

9 The ice breaker question uses an example from Amartya Sen's book *The Idea of Justice*. In this example three children are fighting over a flute. Each of the children makes a claim on the flute, but they all appeal to a different principle of justice (Sen 2009). At the very beginning of the questionnaire, respondents were presented with this scenario and asked to decide which of the three children, Anne, Bob, or Carla, should get the flute.