

The Factorial Survey Experiment on "Distribution of Work in Partnerships" in the German Family Panel (pairfam)

**Content and Technical Information on the
Partner Vignettes**

Sabine Düval, Katrin Auspurg

Department of Sociology

Ludwig-Maximilians-Universität, München

May 2022

Funded as long-term project by the German Research Foundation (DFG)

Cite as:

Düval, Sabine, and Katrin Auspurg (2022): The Factorial Survey Experiment on "Distribution of Work in Partnerships" in the German Family Panel (pairfam) - Content and Technical Information on the Partner Vignettes. pairfam Technical Paper No. 20. <https://doi.org/10.5282/ubm/epub.92002>

I. Introduction

A factorial survey experiment was implemented in the 10th wave (2017/18) of the German Family Panel (pairfam).¹ In the social sciences, factorial survey experiments (also known as vignette experiments) are widely used to measure respondents' attitudes, beliefs, and/or behavioral intentions. In short, respondents are asked to evaluate one or several scenarios (i.e., vignettes) describing hypothetical situations, objects, or individuals based on a multifactorial experimental design. In the experiment implemented in the pairfam study, the vignettes were used to describe hypothetical couples with different family statuses and work arrangements. Respondents were asked to indicate whether, in their opinion, the described workload of one of the two partners (either the male or female partner) was appropriate, or whether they think it should be lower or higher. Within the descriptions of these couples' situations, selected characteristics (i.e., dimensions) were experimentally varied in their values (i.e., levels), enabling an estimation of the impact of the experimental variations on respondents' evaluations of appropriate workload. As these dimensions are varied independently from each other, this method allows for a more precise measurement of judgment principles and normative attitudes than do single-item questions. In a vignette experiment, respondents are required to make multidimensional choices; in other words, weigh different dimensions against each other. In this way, the relative importance as well as possible interactions or trade-offs between the experimentally varied dimensions can be determined. At the same time, these designs have the potential to reduce biases due to social desirability, as respondents have multiple reasons options to justify a particular vignette or evaluation (Auspurg and Hinz 2015; Hainmueller et al. 2014). For further information on factorial survey experiments, see Auspurg and Hinz (2015), Mutz (2011), and Wallander (2009).

The factorial survey experiment implemented in the 10th wave of the pairfam study was used to collect evaluations of the appropriateness of one of the partner's workload as described for the hypothetical couple. More precisely, respondents were to rate the appropriateness of the *division of housework* or the *total workload* for the hypothetical male or female partner described in the vignette. This evaluation task and the requested evaluation of the man's vs. the woman's share of work was randomly

¹ All analyses in this Technical Paper are based on data from the 10th wave of the German Family Panel (pairfam), Release 10.0 (Brüderl et al. 2019). A detailed description of the study can be found in Huinink et al. (2011).

assigned using a between-respondent experimental design. All evaluations relied on eleven-point response scales. The following dimensions were experimentally varied across the three vignettes and evaluated by each respondent: partners' time in paid and unpaid work, both partners' financial resources, and family status (i.e., marital status, presence and age of children). All hypothetical vignette couples were described in heterosexual partnerships, and the main focus of the design was to observe whether the gender of the vignette person impacts the evaluation: Should female vignette persons do more (house-) work than male vignette persons? Did this hold when both partners had the same paid working hours and/or earnings? To what extent do respondents believe that male or female vignette persons are entitled to "buy" their way out of housework by doing higher shares of paid work? All in all, this specific vignette design was explicitly drafted to disentangle gender role theories from alternative theories aiming to explain gendered work distributions within couples (see Düval 2022; Düval and Auspurg 2018).

In addition, the vignette experiment was designed to contribute to methodological research and to test theories on the conditionality of norms. The amount of information on the hypothetical couple's characteristics was therefore also varied in between-respondent splits (e.g., some vignettes contained information on the exact working hours and earnings of both partners, while other vignettes, in other survey splits, did not). This was done to explore whether respondents assume gender-typical work arrangements (i.e., the female partner does less paid work/has lower earnings than the male partner) when exact information on these characteristics is missing, as is the case in standard survey items exploring gender role attitudes in many surveys.

A detailed description of all 10 between-respondent splits and different evaluation tasks can be found in Düval and Auspurg (2020). This technical report includes general information on motivations and the background of the experiment, as well as a detailed description of the factorial survey experiment, the experimental design, and the resulting vignette data set (i.e., "anchor10_vig.dta" or "anchor10_vig.sav"). In addition, information on the realized anchor sample and first descriptive results are included. Note that Düval and Auspurg (2020) describe the vignette module implemented in the wave 10 anchor questionnaire. The additional vignette module implemented in the wave 10 partner questionnaire is introduced in this technical paper. In the following, core information on the vignette design used in both the anchor and

partner survey is described. The main purpose of this report is to document the (few) differences across the vignette modules implemented in the anchor vs. the partner survey, for example regarding survey mode.

Figure 1. Example vignette (variable dimensions underlined; full information condition; evaluation of total workload)

A married couple has an 8-year-old child. She works 30 hours per week, he works 20 hours per week.
 Her contribution to their monthly household income is approximately half of his.
 Both are normally responsible for 50% (15 hours per week) of the weekly housework (e.g., laundry, cooking, cleaning, repairs).
 She is responsible for a smaller share of the child care than is he.

How appropriate do you think her share of the total workload (incl. housework, paid employment, and child care if applicable) [answer format 2: her share of the housework] is?

Her share of the total workload...

should						is					should be
be much						appro-					much lar-
smaller						priate					ger
-5	-4	-3	-2	-1	0	1	2	3	4	5	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

No answer 99

Source: pairfam Group (2019).

II. Description of the Factorial Survey Experiment

Experimental design

Figure 1 shows an example vignette, and Table 1 gives an overview of all dimensions and levels. Anchors and their partners were asked to evaluate an identical set of three vignettes; the experimental setup underlying the construction of vignettes was the same for both anchors and partners. Each anchor-partner couple was part of the same between-respondent split, meaning they received vignettes with the same gender of the vignette person, the same rating tasks, and as part of the same information condition (i.e., providing more or less information on paid work and earnings). The order in which the anchor-partner couple received their vignettes was also the same within each couple. This anchor-partner couple design, holding all experimental treatments constant for both partners within the couple, allows for a straightforward comparison of anchors' and partners' evaluations. Possible differences in evaluations

are then not caused by differences in the experimental stimuli, but rather reveal differences in the two partners' opinions on the appropriateness of different work divisions. However, the survey mode for anchors and partners was different (more information on the survey mode at the end of this section).

Düval and Auspurg (2020) provide more information on the experimental design of the vignettes, including information on the full factorial (vignette) universe that defines all possible vignettes given the experimental variations, the *D*-efficient sampling of the vignettes used in the survey, and the allocation of the anchor respondents to the 10 experimental splits and vignette decks containing three vignettes each.² In short, the *D*-efficient sampling ensured that in the vignettes fraction used in the survey there occurred no correlations between the vignette dimensions and their two-way and three-way interactions (apart from those expected by vignette design; see Düval and Auspurg 2020) occurred. It also ensured that the selection of dimension levels and the combination of these levels were not correlated within the between-respondent splits.

Table 1. Vignette implementation: Overview of all dimensions and levels

		Dimensions	Levels
Exp. split for low-info condition	Exp. split for medium-info condition	1 Marital status	Unmarried/ married
		2 Presence and age of children	No children/ 2 years/ 8 years
		3 Share of child care (man/woman) ^a	Larger share than partner/ smaller share than partner/ same share as partner/ no information
	Exp. split for high-info condition	4 Share of housework (man/woman) ^a	70% (21 hours per week)/ 60% (18 hours)/ 50% (15 hours)/ 40% (12 hours)/ 30% (9 hours)
		5 Hours of paid work per week (man)	40/ 30/ 20 hours/ no information ^b
		6 Hours of paid work per week (woman)	40/ 30/ 20 hours/ no information ^b
		7 Relative earnings (man/woman) ^a	Twice as much as woman/ half as much as woman/ same as woman/ no information ^b

Notes: ^a See information on gender splits: Approximately half of the respondents were informed and asked about the relative share of work of the male/female vignette person. ^b This was varied in between-respondent splits only, see explanation on the three different information conditions in Düval and Auspurg (2020).

Note that the experimental setup was designed for the 4,500 anchor respondents expected to participate in wave 10 of the pairfam survey. Not every respondent was in

² For an overview of the 10 splits, see Table 2 in Düval and Auspurg (2020).

a relationship at that time, and not every possible partner participated in the partner survey or in the vignette module of the partner survey. This reduced the possible number of partners and, thus, the available vignette data. It is therefore important to confirm that the vignettes answered in the partner survey reflect desirable features of the experimental design, such as no confounding of different experimental factors (dimensions). Preliminary information on data quality checks is described in Section 3.

Implementation into the partner survey and survey mode

While the pairfam anchor respondents answered the vignette module in the CASI (computer-assisted self-interview) section of the questionnaire, anchors' partners were asked to evaluate the same three vignettes in a PAPI (paper-assisted self-interview) questionnaire. The partner vignette module was attached as an extra, single sheet of paper to the standard partner questionnaire. As anchors and partners form an anchor-partner couple designed to evaluate the same selection (i.e., deck) of three vignettes in the same experimental split, one unique vignette sheet was prepared for each partner that mirrored the vignettes and experimental split the respective anchor respondent was assigned by random allocation (see Düval and Auspurg 2020 for details).

For interviewers to identify which vignette sheet should be delivered to which respondent, the sheet included list numbers, consecutive numbers, and household numbers allowing for a one-to-one identification of the partner and his/her corresponding vignette sheet. Interviewers were asked to be mindful of the one-to-one identification when distributing the questionnaires. However, it cannot be ruled out that, by mistake, some partner respondents did not receive the correct vignette sheet. The list numbers identifying the vignettes have been checked as to whether they match between anchor and partner; the few cases from the data where this was not the case were removed. When running analyses with the anchor vignettes, we recommend additional checks as described in Section IV.

III. Partner Module: Realized Sample, Data Quality, and Preliminary Descriptive Results

A total of 1,799 partners of the pairfam anchors participated in wave 10 of the pairfam survey. Of these partners, 1,426 (79.3%)³ evaluated at least one of the three vignettes presented as part of the PAPI questionnaire. In light of survey research showing low response rates, this appears to be acceptable, especially when considering that partners did not receive any additional incentive for their participation in the vignette module, and that this module was provided on an extra sheet.⁴ In total, 4,239 vignettes were evaluated by the partners.

Table 2 shows the number of partner respondents and vignette evaluations per experimental split, as well as separately per response task (i.e., evaluations of housework only vs. total workload).⁵

Table 2. Partner vignette module: Number of respondents and evaluations per experimental split and response task

Split	Per Split		Per Response Task		
	Respondents	Vignette Evaluations	Respondents	Vignette Evaluations	
1	56	167	324	963	Response task: housework
2	51	148			
3	55	164			
4	62	185			
5	50	150			
6	50	149			
7	134	399	958	2,848	Response task: total workload
8	144	427			
9	350	1,044			
10	330	978			
Σ	1,282	3,811	1,282	3,811	

³ As noted in Section II, not all list numbers identifying the vignettes match between anchor and partner respondents. Removing these cases from the data reduces the sample to 1,282 partner respondents and a total number of 3,811 vignette evaluations. Note that all following numbers are based on this corrected sample.

⁴ This step was necessary as each partner/household had to be assigned its own unique sheet based on the vignettes/experimental split to which the anchors were randomly assigned.

⁵ The number of cases for each between-respondent split can be calculated by comparing information provided in Table 2 with information in Table 3 in Düval and Auspurg (2020). For example, for vignette evaluations with the medium-information condition, the values for splits 3, 4, 7, and 8 must be combined.

Note that partners received an identical set of three vignettes as the corresponding anchor respondent – strictly speaking, this implies that partners were not randomly assigned to the vignettes.⁶ However, we can assume that (not) having a participating partner in the partner survey did not influence the allocation of anchor respondents to the vignette decks and splits. Hence, the allocation of anchor-partner couples to the vignette decks can also be understood as (quasi-) random. Indeed, at first glance, it appears as if partner respondents were relatively evenly distributed across experimental splits (see Table 2). Some splits as well as response tasks (i.e., total workload) were purposefully oversampled (see Table 2 in Düval and Auspurg 2020)⁷; therefore, it is expected that, for example, roughly twice as many respondents were allocated to split 7 than split 1. Considering this oversampling, there does not appear to be any systematic non-response related to the different splits or evaluation tasks. The distribution of respondents and vignette evaluations was very similar to the planned distribution of respondents across splits and response tasks, oversampling splits 7, 8, and particularly 9 and 10 with more information on the hypothetical vignette couple, and oversampling the split in which respondents were to evaluate the total workload.

Ideally, the final sample is balanced in terms of level frequency (i.e., all levels of the each dimension occur with about the same frequency). In addition, the vignette dimensions (and their interactions) in the final sample should ideally not be confounded. Both features in unison increase the statistical power to estimate the impact of each dimension on the vignette evaluation. The absence of confounding factors would additionally mean that one could use simple bivariate analyses that do not control for other (unconfounded) vignette dimensions. Although these desirable characteristics of multifactorial experimental designs were optimized by the *D*-efficient sampling of vignettes to be used in the anchor survey, (selective) non-response might have affected these characteristics in the partner survey. As the partner sample is relatively small and a large number of vignette dimensions was used with an experimental split, some level imbalance or correlations might have occurred by

⁶ Not every anchor was in a relationship in wave 10 and not every partner participated in the partner survey or in the vignette module of the partner survey.

⁷ In short, splits 7 and 8 were oversampled by a factor of 2; splits 9 and 10 by a factor of 4.

chance. In this case, it is recommended to use multivariate analyses to analyze the vignette data.

Table A1 in the Appendix shows that the partner sample is quite balanced in terms of level frequency. Overall, the experimental dimensions are not correlated (see Table A2 in the Appendix).⁸ For most analyses focusing on gender role theories and alternative theories to explain gendered work distributions within couples, the absence of any meaningful correlations between the vignette person's gender (i.e., the person whose share of work should be rated by the respondents) and his/her relative earnings, paid working hours, and share of housework is particularly important. In the partner sample, the Spearman correlations between vignette person gender and his/her relative earnings and share of housework were not statistically significant.⁹ However, there was a small but statistically significant correlation [$r_s(\text{sex}, \text{jobP1})=.036$ ($p=.027$)] between the gender of the vignette person and his/her paid working hours.¹⁰ This small correlation should be taken into account when analyzing the partner vignette data. Instead of trusting simple bivariate analyses to contrast evaluations of male versus female vignette persons (as they are sometimes used in unconfounded experiments), we recommend multivariate analyses that control for the other vignette dimensions, thus allowing for the removal of the small bias that might otherwise occur.¹¹

Table 3. Partner module: Preliminary descriptive results

Response Task	N Evaluations	Min./Max.	Mean	SD	Median	25/ 75%-Percentile
Total	3,811	0/10	5.01	2.23	5	4/6
Housework only	963	0/10	4.93	2.31	5	3/6
Total workload	2,848	0/10	5.04	2.21	5	4/6

Notes: All descriptive statistics are at the vignette level.

⁸ Between men's share of housework and women's hours of paid work, there was a statistically significant Spearman correlation ($p=.050$), which was not expected as per experimental design. However, this correlation was rather small ($r_s=.032$). When looking at Pearson's r , the correlation was no longer statistically significant ($r=.030$; $p=.067$). Nevertheless, this correlation is important to consider when analyzing the partner data.

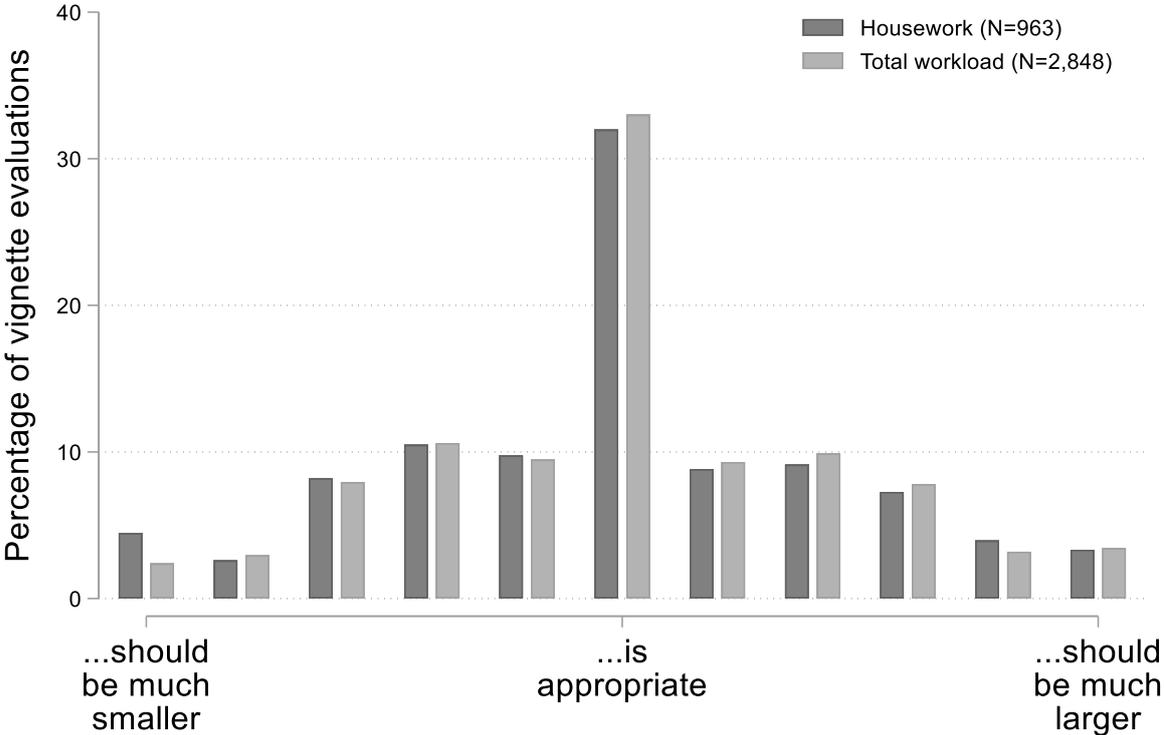
⁹ The corresponding correlations are $r_s(\text{sex}, \text{inc})=.021$ ($p=.196$) and $r_s(\text{sex}, \text{house})=-.003$ ($p=.864$).

¹⁰ Recoding the variables in a way that jobM (jobF) equals the men's (women's) paid working hours, regardless of whether respondents evaluated a male or a female vignette person, no statistically significant correlation with the vignette person's gender was found (see Table A2 in the Appendix).

¹¹ Multivariate regressions have the additional advantage that cluster-robust standard errors that correct for the nested data structure can be implemented, where each respondent evaluated several – three – vignettes. These correlations help to obtain correct p -values and confidence intervals (see Auspurg and Hinz 2015).

Table 3 summarizes preliminary descriptive statistics of the obtained evaluations. Partners' evaluations were distributed across the entire response scale¹², with a total mean response of 5.01 (standard deviation [sd]: 2.23). Both response tasks (i.e., "housework only" and "total workload") had similar evaluation means. These results are comparable to the results of the anchor module of the factorial survey experiment (mean: 5.03; sd: 2.07) (see Düval and Auspurg 2020).

Figure 2. Distribution of vignette evaluations by response task



In addition, a look at the variance in evaluations is necessary to check whether the experimental stimuli had an impact on opinions. Varied responses means that respondents varied in their evaluations either because they reacted differently to the experimental stimuli in the vignettes presented to them, and/or because they differed in their general opinion on who should do how much (house-) work. No variance in evaluations might hint to response sets or other indications of low data quality. A first indication of the variance in responses to this vignette module is given by the

¹² Please note that the original eleven-point evaluation scale was re-coded in the pairfam dataset *anchor10.dta/.sav*, now ranging from 0 "His / Her share of the housework / total workload should be much smaller" to 10 "Her / His share of the housework / total workload should be much larger". This new range has been adopted for the partner data accordingly.

percentiles in Table 3 and in particular by Figure 2, which shows the distribution of evaluations by response task (i.e., “housework only” and “total workload”). In both splits, evaluations were distributed across the entire response scale, meaning that respondents used all possible vignette evaluations. Notably, the middle category (“*His / Her share of the housework / total workload is appropriate*”) is the most frequently chosen response category for both splits. Again, results are similar to the results of the anchor vignette module (see Düval and Auspurg 2020).

IV. Partner Module: Description of the Data Sets

To analyze partners’ responses to the factorial survey experiment, two different data sets are needed: 1) a vignette data set that contains information on the experimental setup of the factorial survey experiment (i.e., the vignette data set “anchor10_vig”); and 2) a data set containing partners’ responses to the factorial survey experiment (i.e., the partner vignette data set “partner10_vig”).

The vignette data set

The anchor respondent and his/her partner form an anchor-partner couple and received an identical set of three vignettes. In order to analyze vignette responses from the partners, no idiosyncratic (setup) vignette data set is needed. The vignette data sets “anchor10_vig.dta” (Stata) or “anchor10_vig.sav” (SPSS) include all vignette data in long format. Data in this format are structured so that each data row represents one vignette. Here, three data lines are included for each couple-specific vignette deck. For more information on the vignette data set, see Düval and Auspurg (2020). A table on the structure of the vignette data set is also provided here (Table 5).

The partner vignette data set

The partner vignette data set “partner10_vig.dta” (Stata) or “partner10_vig.sav” (SPSS) is also structured in long format. Each vignette evaluation is thus displayed in its own data row (indicated by the variable `vignr`, running from 1 to 3). The key variables `vig_index1` (ID indicating unique vignette decks) and `vignr` (indicating the order of the three vignettes) can be used to merge the partner vignette data with the vignette data. To merge these two files with the wave 10 anchor and/or wave 10 partner data sets, the latter two must first be transformed

into long format from wide format. For more detailed information on data management and the analysis of vignette data, see Auspurg and Hinz (2015).

The partner vignette data set also includes anchors' `id` and partners' `pid`, which allow for the identification of the respondents. Partners' responses to the factorial survey experiment are stored in variable `pvig`. In an additional item question, partners were asked for their option on whether the couple should decide how to distribute the workload between partners (item `pvig4`).

Note that partner respondents received the vignette module as part of a PAPI questionnaire. The partner vignette experiment was attached as an extra, single sheet of paper to the partner PAPI questionnaire. Interviewers were asked to be mindful of the one-to-one identification when distributing the questionnaires. However, it cannot be ruled out that some respondents did not receive the correct vignette sheet. As already mentioned, all partner respondents whose individual `vig_index1` did not match the relevant anchor respondent's `vig_index1` have been excluded. Nevertheless, it is possible that matching was not successful for some anchor-partner couples. Therefore, partners were to state their gender (`psex_vig`), month of birth (`pdobm_vig`), and year of birth (`pdoby_vig`) on the vignette sheet to allow for an additional check by the data user.

Acknowledgement

This technical paper uses data from the German Family Panel `pairfam`, coordinated by Josef Brüderl, Sonja Drobnič, Karsten Hank, Franz Neyer, and Sabine Walper. `pairfam` is funded as long-term project by the German Research Foundation (DFG).

V. References

Auspurg, Katrin, and Thomas Hinz. 2015. *Factorial Survey Experiments*. Thousand Oaks, CA: Sage.

Brüderl, Josef, Sonja Drobnič, Karsten Hank, Bernhard Nauck, Franz J. Neyer, Sabine Walper, Philipp Alt, Christiane Bozoyan, Petra Buhr, Christine Finn, Madison Garrett, Henriette Greischel, Nicolai Gröpler, Kristin Hajek, Michel Herzig, Bernadette Huyer-May, Rüdiger Lenke, Lara Minkus, Bettina Müller, Timo Peter, Claudia Schmiedeberg, Philipp Schütze, Nina Schumann, Carolin Thönnissen, Martin Wetzel, and Barbara Wilhelm. 2019. *The German Family Panel (pairfam)*. GESIS Data Archive, Cologne. ZA5678 Data file Version 10.0.0, <https://doi.org/10.4232/pairfam.5678.10.0.0>.

- Düval, Sabine. 2022. "Do Men and Women Really Have Different Gender Role Attitudes? Experimental Insight on Gender-Specific Attitudes toward Paid and Unpaid Work in Germany." Working Paper.
- Düval, Sabine, and Katrin Auspurg. 2018. "Gender Ideologies or Simply a Measurement of Respondents' Beliefs about Inequalities". Presentation at Seminar Analytische Soziologie: Theorie und empirische Anwendungen. Venice International University.
- Düval, Sabine, and Katrin Auspurg. 2020. "The Factorial Survey Experiment on 'Distribution of Work in Partnerships' in the German Family Panel (pairfam)." Pairfam Technical Paper 14:1-20.
- Hainmueller, Jens, Daniel J. Hopkins, Teppei Yamamoto. 2014. "Causal Inference in Conjoint Analysis: Understanding Multidimensional Choices via Stated Preference Experiments." *Political Analysis* 22(1): 1-30.
- Huinink, Johannes, Josef Brüderl, Bernhard Nauck, Sabine Walper, Laura Castiglioni, and Michael Feldhaus. 2011. "Panel Analysis of Intimate Relationships and Family Dynamics (pairfam): Conceptual framework and design." *Zeitschrift für Familienforschung - Journal of Family Research* 23:77-101.
- Mutz, Diana C. 2011. *Population-Based Survey Experiments*. Princeton: Princeton University Press.
- Pairfam Group. 2019. Partner Codebook. Wave 10. 2017/18. Release 10.0. GESIS Data Archive, Cologne. ZA5678 Data file Version 10.0.0, <https://doi.org/10.4232/pairfam.5678.10.0.0>.
- Wallander, Lisa. 2009. "25 Years of Factorial Surveys in Sociology: A Review." *Social Science Research* 38(3):505-20.

VI. Appendix

A1. Partner module, experimental design check: Frequency of dimension levels

Dimension	Level	Realized Evaluations	Share (%)
Marital status	unmarried	1,860	48.8
	married	1,951	51.2
Presence and age of children	no children	1,309	34.4
	2 years	1,256	33.0
	8 years	1,246	32.7
Hours of paid work per week (male partner)	40 hours	1,211	31.8
	30 hours	1,161	30.5
	20 hours	1,124	29.5
	no information	315	8.3
Hours of paid work per week (female partner)	40 hours	1,129	29.6
	30 hours	1,167	30.6
	20 hours	1,200	31.5
	no information	315	8.3
Relative earnings (male partner)	twice as much as woman	776	20.4
	half as much as woman	738	19.4
	same as woman	807	21.2
	no information	1,490	39.1
Share of housework (male partner)	70% (21 hours per week)	774	20.3
	60% (18 hours per week)	736	19.3
	50% (15 hours per week)	793	20.8
	40% (12 hours per week)	718	18.8
	30% (9 hours per week)	790	20.7
Share of child care (male partner)	larger share than woman	668	17.5
	smaller share than woman	630	16.5
	same share as woman	618	16.2
	no information	1895	49.7
Gender of vignette person	male vignette person	1,924	50.5
	female vignette person	1,887	49.5

A2. Partner module, experimental design check: Spearman correlations

	Marital status	Presence /age of children	Hours of paid work per week (man)	Hours of paid work per week (woman)	Relative earnings (man)	Share of housework (man)	Share of child care (man)	Gender of vignette person
Marital status	1.0000	-	-	-	-	-	-	-
Presence/age of children	0.0059 (.7176)	1.0000	-	-	-	-	-	-
Hours of paid work per week (man)	0.0138 (.3937)	0.0164 (.3126)	1.0000	-	-	-	-	-
Hours of paid work per week (woman)	0.0255 (.1162)	0.0077 (.6343)	0.2360* (.0000)	1.0000	-	-	-	-
Relative earnings (man)	0.0121 (.4569)	0.0064 (.6906)	0.1568* (.0000)	0.1623* (.0000)	1.0000	-	-	-
Share of housework (man)	0.0139 (.3903)	0.0076 (.6369)	0.0096 (.5549)	0.0318* (.0500)	-0.0152 (.3486)	1.0000	-	-
Share of child care (man)	-0.0012 (.9385)	-0.5905* (.0000)	-0.0135 (.4062)	0.0272 (.0932)	0.0102 (.5281)	-0.0211 (.1935)	1.0000	-
Gender of vignette person	-0.0105 (.5158)	-0.0146 (.3662)	0.0182 (.2602)	0.0080 (.6195)	0.0063 (.6960)	0.0177 (.2736)	0.0070 (.6663)	1.0000

Notes: The Spearman correlation r_s measures linear dependence between two variables. Numbers close to 0 indicate independence between two variables (vignette dimensions). Therefore, small numbers speak for the quality of the realized vignette fraction (i.e., high orthogonality between dimensions). As the share of child care is dependent on the variable *age of children*, a correlation between those two variables is expected. Also, the small correlations between the hours of paid work per week for male and female partner, as well as the relative earnings are expected and can be explained by design: When there is "no information" on the female partner's working hours, information on the male partner's working hours is also missing, and vice versa. Also, when no information on the relative earnings is available, information on the working hours is missing (i.e., low information condition). p -values in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$