

Supplementary Material: Beliefs about Maternal Labour Supply

Design — The Experimental Survey

Teodora Boneva, Marta Golin, Katja Kaufmann and Christopher Rauh

May 27, 2025

1 Details Supplementary Study

In addition to our non-experimental main data collection, we conduct a supplementary survey into which we embed a survey experiment. This appendix provides details on the sampling method, data and survey design.

1.1 Supplementary Study: Data and Survey Design

The goal of this supplementary study is to examine whether providing respondents with truthful information about the returns to mothers working can shift beliefs and labour supply intentions. For this purpose, we collect survey data from a second representative sample of 1,000 German adults without children, aged between 18 and 45. The data collection was carried out between August and September 2023 in collaboration with the same survey company, *Pureprofile*. We use quota-based sampling to ensure the representativeness of our sample along broad regions of residence, gender, and broad educational attainment (see Table C.1).¹

Our information experiment follows the structure presented in Figure C.1. Respondents to our survey were first given details about the study by Nicoletti, Salvanes and Tominey (2023) that examines how maternal labour supply affects child test scores in the context of Norway. This study has three main features that make it suitable for our purposes. First, the authors use population-wide administrative data, which allows the authors to estimate

¹For quota calculation, we classify the German federal states into the following five groups: (1) Mecklenburg-Vorpommern, Brandenburg, Berlin, Sachsen, Sachsen-Anhalt, Thuringen; (2) Schleswig-Holstein, Hamburg, Niedersachsen, Bremen; (3) Baden-Wuerttemberg, Rheinland-Pfalz, Saarland, Hessen; (4) Bayern; (5) North Rhine-Westphalia.

the effect of maternal labour supply on child human capital for the *average family* in Norway, rather than for a specific subgroup of the population. Second, the study neatly identifies the *causal* effect of maternal work hours while the child is 1-5 years old on child outcomes later in life by leveraging an overlapping peer group approach for identification. Moreover, the identification strategy allows the authors to decompose the total effect into the (causal) direct effect stemming from changes in time allocation and the (causal) income effect. Third, the Norwegian context is comparable to the German one, insofar as the counterfactual to a mother’s time tends to be subsidized, formal childcare. We note that there is no comparable study that credibly identifies the average causal effect of maternal labour supply on child development for Germany, and that the Norwegian context is the closest to the German context among those for which causal estimates of the effect of interest are available. When describing the study to our participants, we state openly that the research was conducted in Norway, and we also introduce respondents to the test score system that is used in Norway to assess children. After being introduced to the study, all participants were asked to guess the results of the study. We then randomly assigned approximately half of our respondents to see the information screen that provided information on the actual results of the study.² Subsequently, we measure respondents’ labour supply intentions (either their own or the labour supply intentions for the mother of their child), as well as post-treatment beliefs about how maternal labour supply (when children are young) affects child outcomes at the point of school entry. We now describe the different survey modules in more detail.

Introduction to the study and guess about study results After eliciting background information on respondents’ demographics, participants to our survey were presented with truthful background information about the study by Nicoletti, Salvanes and Tominey (2023). All participants were told that researchers from the Norwegian School of Economics and the University of York used population-wide administrative data from Norway to conduct a study, in which they analysed how maternal work hours (when children are aged 1-5) affect child test scores at age 15.³ We also introduced the test score scale used in Norway for age-15 standardized assessments and provided our participants with information on the average test scores of Norwegian children (which is 64) and the average work hours of Norwegian mothers when children are 1-5 years old (20 hours per week). Subsequently, we asked participants to estimate the average test score of a child whose mother changed her labour supply when the

²Table C.2 shows that the treatment and control group are balanced in terms of background characteristics.

³When introducing the study, we explained that maternal labour supply decisions can affect child outcomes through different channels, including through household income and changes in maternal time allocation.

child is 1-5 years old from 20 hours per week to 30 hours per week. Respondents' guesses were elicited on a 0-106 continuous scale, which corresponds to the test score scale for age 15 examinations in Norway. We incentivized correct answers with a 1 EUR bonus that was awarded to all respondents who correctly guessed the results of the study. After providing their guesses about the results of the study, participants were randomly allocated to either a control group (approx. 50% of the sample) or a treatment group (approx. 50% of the sample).

Information treatment All participants in the treatment group were then shown the actual results of the study. In particular, participants were told that the researchers found that if a mother increases her working hours from 20 to 30 hours per week when the child is 1-5 years old, her child's test score at age 15 increases on average from 64 to 70 points. We additionally explained that this positive effect is largely due to the fact that more income is available to the household if the mother increases her labour supply. Control group participants were not shown any information at this stage. They continued directly with the rest of the study.

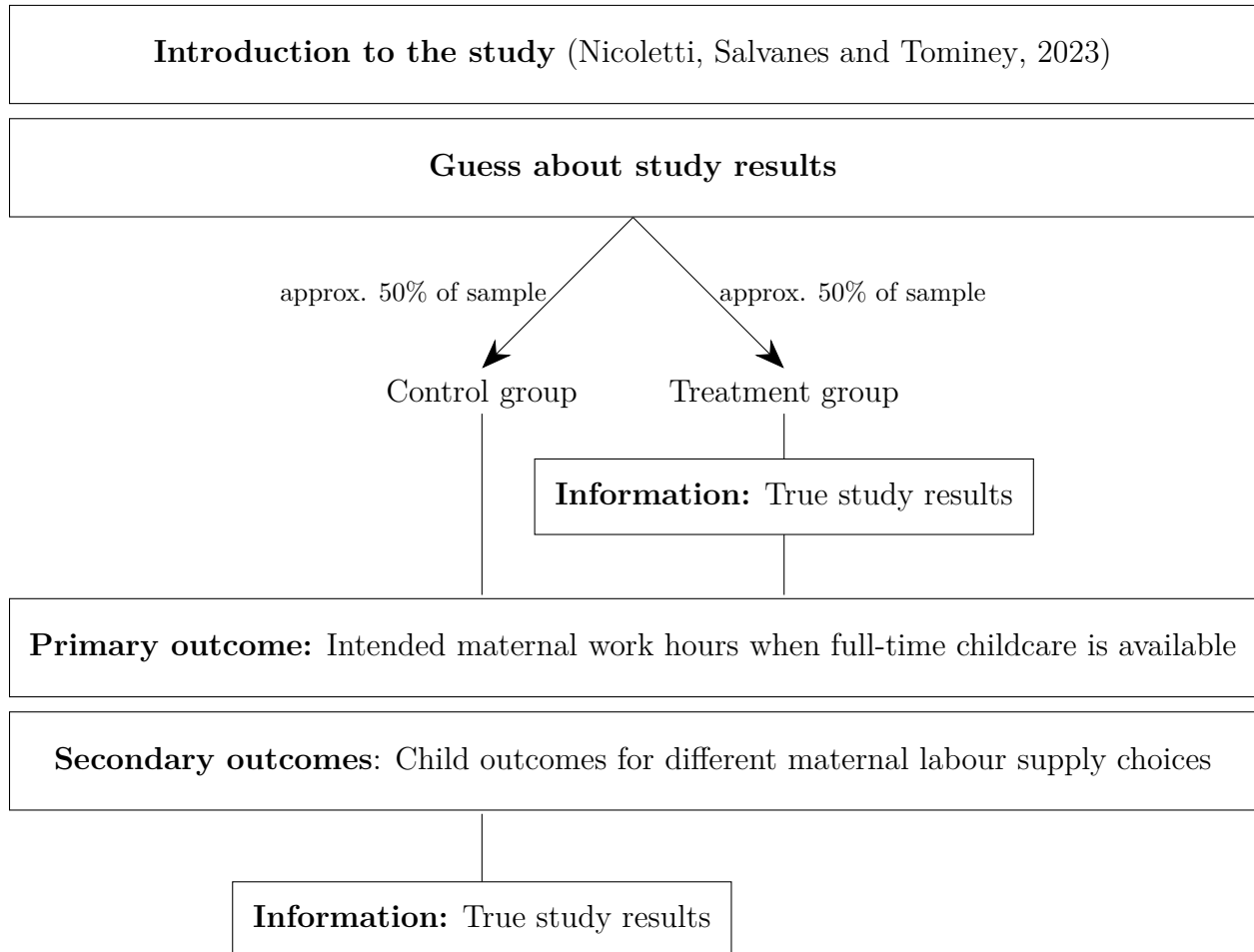
Labour supply intentions We measure labour supply intentions by asking female participants how many hours per week they would most likely work if they had a young child and a full-day place in childcare was available to them. Male respondents are asked a similar question about how many hours per week they think the mother of their child would most likely work, in the same hypothetical situation where they had one child and a full-day place in childcare was available. Labour supply intentions are elicited on a 0-50 continuous scale (hours per week).

Perceived returns to maternal labour supply We then elicited respondents' beliefs about how maternal labour supply affects child outcomes at the time of school entry (at age 6). We measure these beliefs with the same hypothetical scenarios used in our main survey (vignette A).

Demographics After eliciting our outcomes of interest, we measured further background characteristics of the respondents. At the end of the survey, control group participants were shown the same information screen that treated participants saw right after eliciting their guesses. The survey then concludes for all subjects.

1.2 Design Supplementary Study: Figures and Tables

Figure C.1: Structure of the survey experiment



Notes: This figure shows the structure of our experimental design.

Table C.1: Sample representativeness - Supplementary sample

	Sample (%)	National population (%)
Female	42.70	43.01
University degree	25.50	25.51
Region:		
Group 1	18.90	18.87
Group 2	15.90	15.86
Group 3	26.80	26.82
Group 4	14.50	14.50
Group 5	23.90	23.94

Notes: This table displays the share of people between 18 and 45 without children that are residents in each of the five broad regions as well as the share of women and people with a university degree in our sample (column 1) and in the national population (column 2). The national population distribution across regions as well as the share of women and people with a university degree has been calculated from the relevant population of respondents to the German Socio-Economic Panel (GSOEP), using the survey weights provided in the GSOEP.

Table C.2: Balance table - Supplementary sample

Variable	Control	Treatment	Difference
Female	0.411 [0.492]	0.444 [0.497]	0.033 (0.294)
Age in years	33.351 [7.255]	33.462 [7.486]	0.111 (0.813)
East Germany	0.183 [0.387]	0.196 [0.397]	0.013 (0.599)
Married	0.161 [0.368]	0.188 [0.391]	0.027 (0.264)
Work full time	0.603 [0.490]	0.579 [0.494]	-0.025 (0.430)
Own mother worked	0.617 [0.487]	0.625 [0.485]	0.008 (0.796)
Income	36229.840 [25849.205]	35491.805 [24607.055]	-738.035 (0.647)
Observations	504	496	1,000

Notes: The first two columns show the mean and standard deviations of respondents' background characteristics, separately for the control group and treatment group. Standard deviations are reported in square brackets. The last column shows differences in means between the control group and the treatment group. P-values for a test of differences in means between two groups are reported in parentheses.

References

Nicoletti, Cheti, Kjell G Salvanes, and Emma Tominey. 2023. "Mothers working

during preschool years and child skills: does income compensate?” *Journal of Labor Economics*, 41(2): 389–429.