Institut für Volkswirtschaftslehre

Universität Augsburg

Volkswirtschaftliche Diskussionsreihe

Early Child Care, Maternal Labor Supply, and Gender Equality: A Randomized Controlled Trial*

Henning Hermes, Marina Krauß, Philipp Lergetporer, Frauke Peter, Simon Wiederhold

Beitrag Nr. 345, April 2024

Early Child Care, Maternal Labor Supply, and Gender Equality: A Randomized Controlled Trial*

Henning Hermes[a], Marina Krauß[b], Philipp Lergetporer[c], Frauke Peter[d], Simon Wiederhold[e]

April 9, 2024

[a] *ifo Institute Munich*
[b] *University of Augsburg*
[c] *Technical University of Munich, TUM School of Management & ifo Institute Munich*
[d] *German Centre for Higher Education Research and Science Studies (DZHW)*
[e] *Halle Institute for Economic Research (IWH), MLU Halle-Wittenberg & ifo Institute Munich*

Abstract

We provide experimental evidence that enabling access to universal early child care increases maternal labor supply and promotes gender equality among families with lower socioeconomic status (SES). Our intervention offers information and customized help with child care applications, leading to a boost in child care enrollment among lower-SES families. 18 months after the intervention, we find substantial increases in maternal full-time employment (+160%), maternal earnings (+22%), and household income (+10%). Intriguingly, the positive employment effects are not only driven by extended hours at child care centers, but also by an increase in care hours by fathers. Gender equality also benefits more broadly from better access to child care: The treatment improves a gender equality index that combines information on intra-household division of working hours, care hours, and earnings by 40% of a standard deviation, with significant increases in each dimension. For higher-SES families, we consistently observe negligible, insignificant treatment effects.

Keywords: child care, maternal employment, gender equality, randomized controlled trial

JEL: J13, J18, J22, C93

*We thank Kjetil Bjorvatn, Alexander W. Cappelen, Sarah Cattan, Christina Felfe, Luise Görges, Andreas Lichter, Katrine Løken, Markus Nagler, C. Katharina Spiess, Bertil Tungodden, Andrea Weber, Felix Weinhardt, and David Yanagizawa-Drott for helpful suggestions and comments. We received valuable feedback from seminar audiences at the Berlin School of Economics, DICE HHU Düsseldorf, ifo Institute Munich, KU Eichstätt-Ingolstadt, KU Leuven, TUM School of Management, University of Augsburg, University of Mainz, University of Bologna, Berlin School of Economics (BAMS) and Lund University, as well as from audiences at ESA 2021 and 2022, RES 2022, SOLE 2022, AFE 2022, ESPE 2022, Economics of Education Field Committee of the German Economic Association (VfS) 2022, M-BEES Maastricht 2022, EALE 2022, CESifo Conference on Economics of Education 2022, FAIR Midway Conference 2022, IZA Workshop on the Economics of Education 2022, the Ludwig Erhard ifo Conference on Institutional Economics 2023, IWAEE 2023, EEA 2023, VfS Annual Conference 2023, and CESifo Workshop on the Economics of Children. We thank our team of research assistants, interviewers, and backoffice staff for excellent work. We gratefully acknowledge funding from the Jacobs Foundation, the ifo Institute Munich, and support from our home institutions as well as the Johannes Gutenberg University of Mainz. Henning Hermes acknowledges financial support by the Norwegian Research Council through its Centers of Excellence Scheme, project 262675. The study was preregistered in the AEA RCT Registry (AEARCTR-0003181). IRB approval was obtained from the Joint Ethics Committee of the Goethe University Frankfurt and the Johannes Gutenberg University Mainz.

1. Introduction

Despite a considerable convergence during the past decades, gender gaps in labor market outcomes still persist in many countries. A key driver of these remaining gender inequalities is childbirth and especially the unequal division of child care that comes with it (see, e.g., Kuziemko et al., 2018; Andresen and Nix, 2022a,b). One important policy instrument aimed at mitigating the adverse effects of child care duties on maternal labor supply is public provision of universal child care. Past expansions of universal child care have indeed increased maternal labor supply in various countries, including Germany (Bauernschuster and Schlotter, 2015) and Norway (Andresen and Havnes, 2019). However, the benefits of (expanding) universal early child care differ strongly by families' socioeconomic status (SES). Lower-SES families are generally less able to access and utilize universally available education programs, so they benefit less from expanding child care offers (see, e.g., Jessen et al., 2020; Heckman and Landersø, 2022). This unequal access to child care is problematic from an equality perspective, as lower-SES mothers face particularly large gender gaps on the labor market.[1] The fact that lower-SES families react only little, if at all, to expansions in child care also implies that our understanding of the causal impact of child care access on the labor supply of lower-SES mothers is very limited.

We study the causal impact of improving access to child care for lower-SES mothers in Germany. The country provides an ideal setting for investigating how child care access may reduce gender gaps in the labor market for several reasons: First, gender gaps in the labor market are very high compared to other industrialized countries (OECD, 2023). Second, the negative impact of children on mothers' labor market outcomes is particularly strong (Kleven et al., 2019). Third, even mothers who return to the labor market after childbirth rarely do so full-time (Ilieva and Wrohlich, 2022). At the same time, Germany provides universal early child care with regulated quality standards and low child care fees due to heavy public subsidization (see, e.g., Felfe and Lalive, 2018). However, parental demand for slots exceeds supply, and lower-SES families are strongly underrepresented in early child care.[2]

[1] The average gender employment gap for lower-SES mothers in OECD countries is 33.1 percentage points (pp), nearly twice as large as the average gap for higher-SES mothers (17.3 pp) (OECD, 2017b).

[2] The difference in early child care enrollment between parents with and without a college entrance qualification ("Abitur") is 14 pp. Differences in the demand for child care account for only about half of this gap (Jessen et al., 2020). We use the same main definition of SES throughout the paper.

To overcome the challenge that child care enrollment of lower-SES families is often unresponsive to child care expansions, we conduct a randomized controlled trial (RCT) in a sample of $n = 607$ families with a treatment designed to specifically increase enrollment of lower-SES families. In particular, to enable child care access for lower-SES families, we provide all families in the treatment group with information and customized assistance in the child care application process.

Nine months after the intervention, lower-SES families in the treatment group are significantly more likely to be enrolled in child care (also see Hermes et al., 2021). Specifically, their likelihood of securing a full-time slot increases by 10.4 pp, corresponding to 78% of the control group mean, and they use on average four more hours of child care per week. Generally, across enrollment outcomes, the treatment closes about half of the SES gap observed in the control group.

By enabling access to early child care, our intervention strongly affects the labor supply of lower-SES mothers measured 18 months after the intervention (when children are 2–3 years old). The treatment increases mothers' probability of working full-time by 9.2 pp, which represents an increase of more than 160% relative to the control group mean.[3] Again, this closes about half of the SES gap. Net household income in the treatment group increases by 10%, suggesting that the increase in maternal full-time employment does not lead to a corresponding decrease in fathers' labor supply. Among working mothers, the treatment increases earnings by 22%.[4] Reassuringly, the intervention has no discernible impact on labor market outcomes for higher-SES mothers, which is consistent with the lack of treatment effects on child care enrollment for higher-SES families.

We also study potential mechanisms underlying the treatment effect on full-time employment for lower-SES mothers. The treatment increases the weekly care hours at child care centers by 4.0 hours and paternal care by 2.6 hours, while simultaneously reducing reliance on alternative arrangements like grandparental care by 2.3 hours per week. Thus, the treatment affects how families combine different forms of non-maternal care, leading to a more equal division of child care hours between mothers and fathers. A mediation analysis reveals that 56% of the overall treatment effect on full-time employment can be attributed to these shifts in non-maternal care, with care hours provided by child care centers (29%) and fathers (20%) being most relevant.

[3] We follow the standard OECD definition for full-time work of 30 or more hours per week.

[4] Additional analyses reveal that the full-time employment effect for lower-SES mothers is due to increased working hours (intensive margin) rather than increased labor market participation (extensive margin) (see Section 6.2).

Finally, we analyze how improved access to universal child care affects gender equality in lower-SES families. We construct an index of intra-household gender equality that combines household-level information on the division of paid and unpaid work, and on the earnings gap between mothers and fathers. The treatment improves this gender equality index by as much as 40% of a standard deviation. Each of the three dimensions of the gender equality index improves significantly due to the treatment. Specifically, we observe increases in gender equality of 20% in paid work, 30% in care hours, and 32% in earnings. The treatment fully closes the control-group SES gaps in these gender equality outcomes. In fact, gender equality in treated lower-SES families closely resembles the *distribution* of gender equality in higher-SES families in the control group. These findings strongly emphasize the importance of access to universal child care for moving towards a more gender-equal society.

We verify our results in several robustness checks. First, we show that results are robust to accounting for attrition (which is generally low and independent of the treatment status) using inverse probability weighting and bounding analysis. Second, drawing on birth registry data, we show that our study participants closely mirror the full cohort of families with children below one year of age. Accordingly, our results do not change when directly accounting for selection into our study by using propensity score weights. Treatment effects are also robust to correcting for multiple hypothesis testing and when using randomization inference. Finally, while our main SES definition is based only on maternal education, we show that effects are qualitatively similar for alternative SES definitions that also encompass pre-treatment household income and single mother status.

Our paper presents the first evidence from an RCT that enabling lower-SES families to access universal early child care boosts maternal labor market participation and improves intra-household gender equality. We thereby contribute to several strands of literature.

First, previous research in quasi-experimental settings has provided compelling evidence on the effects of child care on maternal labor supply.[5] But evidence on how child care affects *lower-SES* families — who are characterized by low child care participation

[5] While quasi-experimental studies focusing on child care for older children (aged 3–6 years) find mixed overall effects on maternal labor market outcomes (see, e.g., Baker et al., 2008; Cascio, 2009; Fitzpatrick, 2010; Havnes and Mogstad, 2011; Bauernschuster and Schlotter, 2015), those focusing on child care for younger children — as we do — find mostly positive effects (e.g., France: Goux and Maurin (2010); Switzerland: Ravazzini (2018) ; Italy: Carta and Rizzica (2018); Belgium: Dujardin et al. (2018); Norway: Andresen and Havnes (2019); Andresen and Nix (2022a); Germany: Müller and Wrohlich (2020); Huebener et al. (2020)). Several papers show that effects on maternal employment are especially pronounced at the intensive margin (i.e., full-time employment rates or working hours; see Ravazzini, 2018; Andresen and Havnes, 2019; Huebener et al., 2020).

and large gender gaps in labor market outcomes (see, e.g., Van Lancker and Ghysels, 2016; OECD, 2019; Pora and Wilner, 2019; Cascio, 2021; Flood et al., 2022) — has been scarce.[6] This is due to the fact that child care enrollment of these families is often barely, if at all, affected by child care expansions or reforms typically exploited in quasi-experimental studies (see, e.g., Van Lancker, 2018; Scholz et al., 2018). Consequently, previous studies have hardly been able to identify effects of child care on labor market outcomes of lower-SES mothers. We overcome this identification challenge by designing a randomized intervention that specifically targets child care participation of lower-SES families.[7]

Very recently, a few papers have experimentally investigated how access to child care affects maternal labor supply.[8] Our paper differs from these studies in at least two important dimensions: First, we focus on *early* child care for children under the age of three, which is arguably the most critical period for mothers to successfully return to the labor market after child birth (OECD, 2011). Second, we focus on the effect of *universal* child care, as opposed to targeted programs serving only a small subset of children. In many countries, universal child care programs are available nationwide, highlighting the potential scalability of our results. Closest to our study, Attanasio et al. (2022) analyze a lottery conducted by the government of Rio de Janeiro among families who applied for early public child care. The authors find no effects of winning the lottery on the labor supply of parents, but show an increase in the labor supply of grandparents living with the child (which is common in the Brazilian context). Importantly, in contrast to Attanasio et al. (2022), our study also reaches families who would not have applied for child care without the intervention. Our results complement their findings, suggesting that existing institutional differences (related to, e.g., the child care system, labor market institutions, and household organization) limit the transferability of findings between developing and developed country contexts.

[6]Neither the lower use of child care nor the lower maternal labor market participation in lower-SES families appear to merely reflect their preferences. Compared to their higher-SES counterparts, lower-SES families are more likely to have an unmet demand for child care (Jessen et al., 2020) and lower-SES mothers are less likely to work as much as they desire (see, e.g., Harnisch et al., 2018; Geis-Thöne, 2021).

[7]Intriguingly, our finding that enabling access to early child care has particularly strong and positive effects on lower-SES mothers' full-time employment is consistent with recent theoretical equilibrium models on the impact of increasing access to early child care (Borowsky et al., 2022).

[8]Wikle and Wilson (2022) reanalyzed the experimental impact study of Head Start — a non-universal child care program serving about 3% of children in the U.S. — and find a marginally significant positive effect on maternal full-time employment (see also Schiman, 2022). Evidence from developing countries is more mixed: While Clark et al. (2019) and Ajayi et al. (2022) find positive effects of child care access on maternal labor market outcomes in Kenya and Burkina Faso, respectively, Attanasio et al. (2022) and Bjorvatn et al. (2022) find no effects in Brazil and Uganda, respectively.

Finally, our study adds to a large literature investigating gender inequality within households and, more generally, gender gaps in the labor market (see, e.g., Goldin, 2021, 2023; Lundberg, 2023). There is ample international evidence that the existing gender gaps in the labor market are due to the fact that parenthood has a much larger negative effect on labor market outcomes of mothers compared to fathers, often referred to as "child penalty" (see, e.g., Kuziemko et al., 2018; Kleven et al., 2019).[9] Andresen and Nix (2022b) investigate the mechanisms behind the "child penalty" for mothers, emphasizing the unequal division of child care duties to the disadvantage of mothers. Our RCT addresses this very issue, showing that experimentally increasing non-maternal care hours has a strong positive effect on maternal labor market outcomes and intra-household gender equality.

2. Institutional Setting: Maternal Labor Supply and Child Care in Germany

In Germany, mothers experience some of the highest child penalties in terms of full-time employment and earnings among OECD countries (OECD, 2017a; Kleven et al., 2019). For example, only 17% of mothers but 86% of fathers work full-time when their youngest child is two years old (BPB, 2021). Most mothers do not return to full-time employment even as their children grow older. For instance, when the youngest child is 15–17 years old, only one-third of mothers work full-time, compared to 88% of fathers (see Appendix Figure A1).[10] The long-term drop in maternal earnings after the birth of the first child is as high as 61%, while fathers' earnings are not affected by childbirth (Kleven et al., 2019). Yet, child penalties differ considerably by socioeconomic background. Among mothers with children aged 0–3 years, 14% of lower-SES mothers are employed, compared to 46% of higher-SES mothers. The relative gap in full-time employment is particularly large: Only 4% of lower-SES mothers work full-time, as opposed to 16% among higher-SES mothers (BMFSFJ, 2012, 2020). The low share of lower-SES mothers in full-time employment has substantial negative implications for their career progressions, lifetime earnings, and pensions. In fact, the gender pension gap in Germany is one of the largest in the OECD (OECD, 2017b).

[9] Following the literature, we also use the term "child penalty" for conceptual accuracy. The term is sometimes criticized for its normative connotation (Berniell et al., 2021), and we do not mean to imply any value judgment regarding children or labor market decisions by using the term.

[10] One reason for these different labor market experiences of mothers and fathers is that mothers typically assume greater child care responsibilities, leading to reduced participation in paid work compared to fathers. The share of employed mothers working part-time is exceptionally high in Germany compared to other OECD countries (see, e.g., OECD, 2020; Müller and Wrohlich, 2020; Ilieva and Wrohlich, 2022).

To increase maternal (full-time) employment rates, Germany has implemented and expanded a variety of family-friendly policies in recent years. For instance, Germany provides paid parental leave (*Elterngeld*) after childbirth for 12 months (14 months if each parent takes at least 2 months), offering parents a compensation of approximately 67% of their pre-childbirth income. If both parents work only part-time, this paid leave period can be extended up to 32 months (*Elterngeld Plus*). Independent of taking paid parental leave, parents are entitled to an (unpaid) job-protected leave period of up to three years (*Elternzeit*).[11]

Furthermore, Germany has greatly expanded its universal child care system over the past decade (Education Report, 2020). However, despite the existence of a legal entitlement to a child care slot for all children from their first birthday onward, demand for early child care slots still exceeds supply. At the same time, the quality of child care is generally high, slots in child care are heavily publicly subsidized (on average, families pay only about 250 EUR per month), and low-income families are eligible for fee exemptions. Most child care centers are operated by municipalities or non-profit organizations; only 1% of centers are run by for-profit organizations (Education Report, 2020). Therefore, similar to many other European countries, there is little competition among providers (Spiess, 2008). Although these family-friendly policies have generally led to increases in maternal labor supply, significant differences between mothers and fathers remain, both at the extensive and intensive employment margin (see, e.g., Müller and Wrohlich, 2020; Ilieva and Wrohlich, 2022; Destatis, 2023).

Lower-SES families use child care substantially less often than their higher-SES counterparts.[12] For example, Jessen et al. (2020) show that children of parents without college entrance qualification are 14 pp less likely to be enrolled in early child care than children of parents with college entrance qualification, which translates into a 37% lower enrollment rate.[13] One major reason for this SES gap is the fierce competition for available slots, stemming from substantial rationing (conditional on demand, about one out of three children does not find a child care slot, see Jessen et al., 2020). Problematically, the allocation of child care slots is decentralized, unstructured, and non-transparent, giving

[11]In practice, the take-up of these policies is very unevenly distributed between mothers and fathers: Unpaid parental leave is taken almost exclusively by mothers. For paid parental leave, three out of four recipients are mothers. Even if fathers take paid parental leave, their leave duration is much shorter than for mothers (3.1 months for fathers vs. 14.0 months for mothers, Destatis, 2022).

[12]This is not only the case in Germany, but in many countries around the world (see Flisi et al., 2022; OECD, 2023).

[13]This gap remains substantial even when accounting for SES differences in demand for child care (Jessen et al., 2020).

well-informed and well-organized parents an advantage in securing a slot. Furthermore, lower-SES parents typically have less time, money, and social capital to invest in the child care application process, which reduces their chances of obtaining a slot. As a consequence, lower-SES families are strongly underrepresented in early child care (Jessen et al., 2020) and have benefited less than higher-SES families from past expansion of child care (Cornelissen et al., 2018; Scholz et al., 2018).

We designed an intervention that specifically addresses the problems lower-SES families face when applying for a child care slot. In particular, the intervention provided relevant information and personalized assistance to help families navigate the complex child care application process. We will now first describe recruitment and our sample before returning to a detailed description of how our treatment addressed the challenges described above (see Section 3.3).

3. Study Design

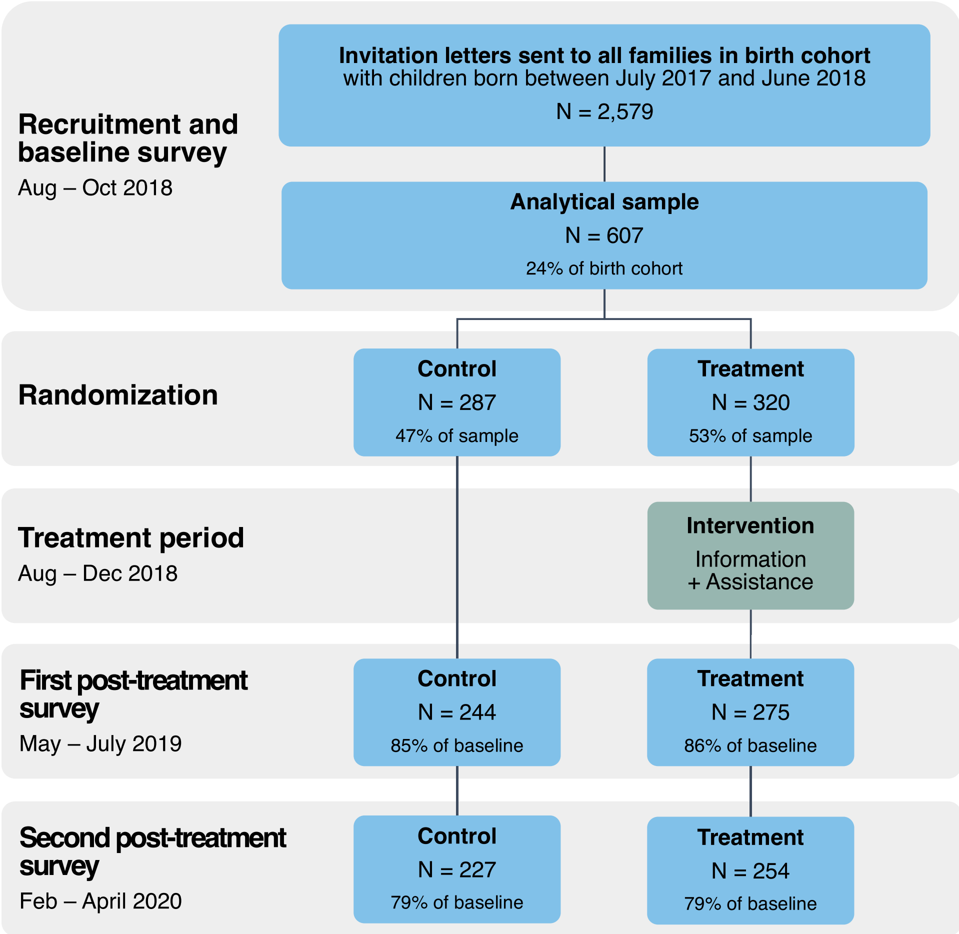
3.1. Recruitment of Sample and Data Collection

The key challenge to conducting our project was to recruit a sufficient number of lower-SES families with very young children, and — perhaps even more challenging — to retain a substantial fraction of these families in the sample over a period of almost two years. To succeed in this, we relied on (i) data from the administrative birth registry comprising the entire birth cohort of two large cities, (ii) an elaborate recruiting strategy, and (iii) a comprehensive set of measures to maximize participation in the post-treatment surveys. Figure 1 provides an overview of the timeline of the study.

Recruitment and Baseline Survey. Our sample consists of families with children aged zero to one year in two large cities in Rhineland-Palatinate (>100'000 inhabitants). In July 2018, we sent postal invitation letters to all 2,579 families with children born between July 1, 2017, and June 30, 2018, using address data from the birth registries of these two cities (to which we were granted access specifically for this study). The letter informed families of the possibility of participating in a research project on 'the life of parents with young children'.[14] In addition to informing parents about the study timeline and the institutions involved, the letter also stated that the study is supported by the Ministry of Education of Rhineland Palatinate and that participation is voluntary. The letter addressed the mother

[14]To obfuscate the exact research question, we referred to the study as "ELFE-Studie" (the acronym stands for Eltern, Leben, Familie, Erziehung, translating to *Parents, Life, Family, and Education*) in our communications with families.

Figure 1: Study Timeline, Sample Sizes, and Attrition



Notes: Figure shows the timeline of the study, as well as sample sizes at each stage.

of the child and announced that a staff member would visit the families at home to conduct the first interview. To optimize communication with the families, we commissioned a public relations and design agency. They created a professional corporate study design tailored to families with young children for all our communication (i.e., a logo, letters, postcards, etc. — see Appendix Figure H1 for examples of the design).

The baseline survey was conducted between August and October 2018 by ten interviewers whom we selected and trained specifically for this study. These interviewers went door-to-door, visiting each family up to three times to encounter them at their home. Families were randomly assigned to one of the interviewers, and each interviewer received an individualized list of families and a random starting point for the recruitment tour in order to achieve a broad geographic coverage. At the first encounter between the interviewer and a parent, in almost all cases the biological mother (94%), the interviewer asked the parent if he or she wanted to participate in the study. If parents agreed to participate and signed the consent form, the interviewer conducted a computer-assisted face-to face interview (CAPI). The median interview time was 23 minutes. To recruit a substantial fraction of parents from lower-SES backgrounds, we paid a generous cash incentive of 20 EUR for participation in the baseline survey (about 24 USD in 2018).

We recruited a total of 607 families at baseline. All families had been pre-randomized into treatment and control groups. To maximize compliance with the treatment, parents in the treatment group were shown a four-minute information video on the interviewer's tablet computer immediately after the interview was completed (see Section 3.3). Up until this point, the interviewers were blind to the treatment condition of the respective family. Relying on the birth registry data, we can show that the sample of families who participated in our study represents the entire cohort of families with children below one year very well. The samples do not differ in terms of area of residence (zip-code level) or single-parent status. We were also able to reach a large share of parents with migration background (41%), albeit still smaller than the share in the entire cohort (51%). As intended, the share of first-time parents in our sample is higher than in the entire cohort.[15] Reassuringly, applying propensity score weights that reflect the probability of participating in our study does not change our results (see Section 6.3 for details). Moreover, and most importantly for the purpose of our study, we succeeded in recruiting a large share of lower-SES families ($> 40\%$).

Post-treatment Surveys. Nine months after the baseline survey, in summer 2019, we conducted our first post-treatment survey via telephone to collect information about child care applications and enrollment ($n = 519$, 86% of the baseline survey). Another nine months later, in spring 2020, we implemented our second post-treatment survey via telephone to elicit information about maternal labor supply ($n = 481$,[16] 93% of the first post-treatment survey and 79% of the baseline survey).[17] The median interview duration of the second post-treatment survey was 44 minutes. Similar to the baseline survey, we

[15]First-time parents were given recruitment priority as we expected them to be especially vulnerable to barriers in the child care application process.

[16]In total, 483 parents (80% of the baseline survey) participated in the second post-treatment survey. We exclude two observations for which we do not have information on maternal employment.

[17]Around 20% of interviews were conducted during the first COVID-19 lockdown, which began in Rhineland-Palatinate on March 16, 2020. Participation after the beginning of the lockdown was not selective with respect to the treatment: When we regress an indicator for participating post-lockdown on our treatment indicator and its interaction with a higher-SES dummy, we do not find any significant differences between treatment and control groups. Furthermore, our estimates are robust to controlling

paid families a 20 EUR participation fee with vouchers for online stores or by bank transfers. Again, almost all interviews (93%) were conducted with the child's biological mother, providing first-hand information on maternal labor supply. The remaining interviews were conducted with the child's biological father, from whom we obtained information about maternal labor supply.[18]

We exerted great effort to find and interview parents for the post-treatment surveys. In between surveys, we regularly sent postcards and additional materials to stay in contact with families in treatment and control groups.[19] We also set up a WhatsApp account to which parents could easily send any updates of their contact information — in particular, mobile numbers changed quite frequently in our sample. When parents could not be reached by phone, we implemented several additional measures to contact them. First, we sent these families electronic and postal reminders. Furthermore, we hired two additional interviewers to personally contact those families who reacted neither to calls nor to the reminders. If these interviewers did not encounter families at their home addresses, we filed registry-data inquiries to re-contact families who had moved. Finally, we sent out a shortened online version of the survey to reach those parents who could not be contacted otherwise ($n = 21$ cases in the second post-treatment survey).[20]

3.2. Sample Description

Sample Characteristics. Table 1 shows the characteristics of our analytical sample, which consists of participants in the second post-treatment survey ($n = 481$). All variables shown in the table were collected at baseline. The upper panel of Table 1 provides pre-birth labor market outcomes of mothers. Almost half of the mothers (49%) were full-time employed before the child was born. 72% of mothers worked part-time or full-time before childbirth. Average net earnings of working mothers amount to 1,777 EUR. Net earnings drop to 1,276 EUR when including non-working mothers with zero earnings. Although such information can be considered highly sensitive, maternal earnings are missing for only 5% of the sample.

for a lockdown indicator (see Appendix Table E2, Column (1)). Thus, we are confident that the lockdown did not impact our results.

[18] Results are robust to excluding all interviews conducted with fathers (see Appendix Table E2, Column (2)). All regressions control for whether the mother or father participated in the survey.

[19] For example, as a "thank you" for participating in the study, we sent a postcard to families that could be converted into a small memory game suitable for the children's age at that time (see Appendix Figure H1).

[20] Our results are robust to excluding these parents or to controlling for the survey mode (result available upon request).

The lower panel of Table 1 reports sociodemographic characteristics of mothers. In 97% of the participating families, the mother is the main caregiver for the child. Mothers' average age is 32 years, 36% of mothers were not born in Germany (i.e., first-generation immigrants), and 13% were working or searching for a job at the time of the baseline survey. The average net household income per month is 3,245 EUR.[21]

Table 1: Analytical Sample: Descriptives and Balancing Tests

	All (1)	Control (2)	Treatment (3)	Δ(3)-(2) (4)	p-val for (4) (5)	p-val by SES (6)	Observations (7)
Pre-birth labor market outcomes							
Mother worked full-time	0.492	0.496	0.488	-0.008	0.868	0.879	474
Mother worked full-time, missing	0.015	0.004	0.024	0.019	0.068	0.629	481
Mother worked	0.717	0.714	0.720	0.007	0.869	0.293	481
Mother's earnings (EUR)	1777	1787	1769	-18	0.867	0.483	328
Mother's earnings, missing	0.318	0.335	0.303	-0.032	0.459	0.410	481
Mother's earnings (0 if not working, in EUR)	1276	1255	1294	39	0.711	0.936	457
Mother's earnings (0 if not working), missing	0.050	0.053	0.047	-0.006	0.779	0.881	481
Sociodemographic characteristics							
Mother is main caregiver	0.971	0.982	0.961	-0.022	0.149	0.261	481
Age of mother (in years)	31.78	31.34	32.18	0.84	0.072	0.117	455
Migration background	0.356	0.358	0.354	-0.003	0.940	0.924	455
Mother works or searches for a job	0.127	0.137	0.118	-0.018	0.546	0.286	481
Household income (EUR)	3245	3135	3345	210	0.208	0.221	461
Household income, missing	0.042	0.035	0.047	0.012	0.508	0.644	481
Interview conducted with the father	0.054	0.040	0.067	0.027	0.181	0.916	481
No school degree	0.029	0.022	0.035	0.013	0.378		481
Lower secondary degree	0.116	0.123	0.110	-0.013	0.656		481
Middle secondary degree	0.243	0.251	0.236	-0.015	0.705		481
College entrance degree	0.611	0.604	0.618	0.015	0.744		481

Notes: Table reports mean values of pre-birth labor market outcomes and sociodemographic characteristics of the mother in our analytical sample. All variables come from the baseline survey (i.e., before the treatment); labor market outcomes refer to one year before childbirth. Column (1) reports mean values for the full sample, Column (2) for the control group, and Column (3) for the treatment group. In Column (4), we show the difference between treatment and control groups, and Column (5) shows the corresponding p-value of a two-sided t-test testing the null hypothesis that values in Columns (2) and (3) are equal. In Column (6), we test whether there are treatment-control differences in the respective variable within SES subgroups. To conduct the test, we regress the variable on the treatment indicator, a higher-SES dummy, and their interaction. Column (6) reports the p-value of an F-test of joint significance of the coefficients on the treatment indicator and its interaction with the higher-SES dummy. *Mother worked full-time, pre-birth* is a dummy equal to one if the mother worked full-time before the child was born, zero otherwise. *Mother worked full-time, missing, pre-birth* indicated the 7 cases in which information on maternal pre-birth work status was not reported. *Mother worked, pre-birth* is a dummy equal to one if the mother worked part-time or full-time pre-birth, zero otherwise. *Mother's earnings, pre-birth* denotes the monthly net earnings of the mother in EUR before the child was born. *Mother's earnings, missing, pre-birth* is a dummy equal to one if either the mother was not working pre-birth (129 cases) or the earnings information for the mother was not provided because the interview was conducted with the father (20 cases) or the mother did not answer the question (4 cases), zero otherwise. In *Mother's earnings, (0 if not working), pre-birth*, we assign zero earnings to mothers who did not work before the child was born; we use this variable when controlling for the pre-birth earnings of mothers. *Mother's earnings (0 if not working), missing, pre-birth* indicates cases in which maternal pre-birth earnings are missing because the father answered the baseline survey or because the mother did not answer the question. *Mother is main caregiver* is a dummy equal to one if the mother is the main caregiver of the child, zero otherwise. *Migration background* is a dummy equal to one if the mother was not born in Germany, zero otherwise. *Mother works or searches for a job* is a dummy equal to one if the mother worked or searched for a job at baseline, zero otherwise. *Household income* is the monthly net household income in EUR. *No school degree*, *Lower secondary degree*, *Middle secondary degree* ("MSA"), and *College entrance qualification* ("Abitur") are dummy variables indicating the mother's highest school degree.

[21] Sample characteristics are very similar to those of mothers with children aged 0–1 years in the German Socio-Economic Panel (SOEP), a nationwide representative household survey (see Goebel et al., 2019; SOEP, 2019). Of these mothers in the SOEP, 65% were working pre-birth with net earnings of 1,580 EUR, and a current net household income of 3,480 EUR.

Definition of SES. As specified in our pre-analysis plan, we are interested in the differential effects of access to universal child care on labor market outcomes for lower- vs. higher-SES mothers. Following previous literature on maternal labor supply (e.g., Dujardin et al., 2018; Müller and Wrohlich, 2020), we define SES based on education, elicited in the baseline survey. In particular, those 61% of mothers ($n = 294$) who have a college entrance qualification ("Abitur") are classified as higher-SES, whereas the 39% of mothers without a college entrance qualification ($n = 187$) are classified as lower-SES (see Table 1). The share of lower-SES mothers in our sample aligns well with the share of lower-SES mothers in Germany overall. In the representative SOEP data, 47% of mothers with children aged 0–1 years are lower-SES according to our classification. Furthermore, our results are robust to using alternative classifications of maternal SES based on mothers' education, household income, and single-mother status (see Section 6.3).

3.3. Treatment

The treatment aimed to reduce potential barriers in the complex child care application process that especially lower-SES parents face (see Section 2 for details) by providing information and personalized application assistance.[22]

The first purpose of the treatment was to address potential gaps in knowledge about the child care application process. Therefore, each family in the treatment group saw a four-minute information video immediately after the baseline survey. The video contained the following information: (i) all parents in Germany have a legal entitlement to child care from child age one onward, (ii) child care in Rhineland-Palatinate is free of charge for all children aged two years and older, and there are fee reductions for younger children (e.g., for lower-income families), and (iii) applying early and to more than one child care center increases the chance of getting a slot.[23] Our treatment was intended to mitigate barriers in the child care application process, not to *persuade* mothers to enroll their child into early child care or to change their work plans. Indeed, the video emphasized that (center-based) child care is only one of several care arrangements, and that it is the families' decision alone how to arrange child care. We consider it highly unlikely that mothers' labor supply preferences are affected by the treatment since maternal labor supply was mentioned neither in the video nor in the second part of the treatment described below. Consistently, our heterogeneity results show that the treatment effect on full-time em-

[22]Based on an ex-ante power analysis, we decided to combine both components into one treatment.
[23]For screenshots of the video and the transcript, see Appendix H.2.

ployment is driven by mothers who reported at baseline that they plan to start working again (see Section 6.1).

In addition to watching the information video, families in the treatment group were offered customized application assistance. The assistance was provided by university students, who received intensive training to support families in achieving their preferred child care arrangement. The assistants' task was to provide personalized support to address the specific problems faced by the families. For instance, they helped families collect information about child care centers and application procedures, assisted with paperwork and filing applications, and reminded parents of important dates, such as open houses at child care centers or application deadlines. Importantly, assistants were not allowed to take over child care duties themselves or to support mothers with job-related tasks (e.g., job search). This second component of the treatment was implemented as an opt-in design, and one-third of families in the treatment group (32% lower-SES and 33% higher-SES) took up the assistance offer. The median time invested by the assistants per family was 1.5 hours, and the median number of contacts with a family (via telephone, email, or in-person) was four. Thus, we consider our intervention to be relatively short and low-cost.

3.4. Randomization, Balancing, and Attrition

We assigned treatment status using stratified randomization (Athey and Imbens, 2017). Using birth registry data, we defined strata based on city of residence (two categories), child's birth quarter (four categories), whether the child lives with both parents (two categories), and first-time parent status (two categories). Within these strata, we randomized families into the treatment group with 50% probability. In our analytical sample ($n = 481$), 227 families (47%) are in the control group and 254 families (53%) are in the treatment group.[24]

Observable characteristics are well balanced between treatment and control groups in our analytical sample (Table 1) as well as at baseline (Appendix Table C1). In fact, none of the differences between treatment and control groups is statistically significant at the 5%-level. Importantly, both groups have very similar shares of mothers without a college entrance qualification (40% and 38%), which we use for defining SES background. Since we estimate treatment effects separately for lower- and higher-SES mothers, we also

[24]These shares are virtually identical in the baseline sample.

verify that observable characteristics are balanced by SES.[25] In sum, our randomization procedure achieved balancing in the full sample and within SES subgroups.

As discussed above, attrition in our study is remarkably low, especially when considering that many participating families have a lower-SES background. About 18 months after the treatment, when we elicit maternal labor market outcomes, we were able to achieve a recontact rate of 79% of the baseline sample (93% of the sample participating in the first post-treatment survey). Results in Appendix Table C2 show that attrition in the second post-treatment survey is not selective with respect to treatment status or pre-birth maternal full-time employment, both in the full sample and in the subsamples of lower- and higher-SES families. In the table, we regress a survey-participation indicator on indicators of treatment status, higher-SES background, and maternal full-time employment before childbirth, as well as on their interactions. All coefficients are close to zero and statistically insignificant. The only exception is that higher-SES mothers not working full-time pre-birth are somewhat more likely to participate in the survey; importantly, however, this does not differ between treatment and control groups.[26]

We can also show that estimated treatment effects are robust to re-weighting the observed data using inverse probabilities of participation in the second post-treatment survey (see Appendix Table C3). Finally, we conduct a bounding analysis as suggested by Lee (2009). Because attrition is very similar in both treatment and control groups, this procedure would only trim one observation, so that our results do not change.

4. Empirical Strategy

We estimate the intention-to-treat (ITT) effects of our intervention on maternal labor market outcomes by ordinary least squares (OLS) using the following regression model:

$$Y_i = \alpha + \beta_1 Treatment_i + \beta_2 Treatment_i \times HigherSES_i + \beta_3 HigherSES_i + \mathbf{X_i'}\delta + \varepsilon_i \quad (1)$$

Y_i is the outcome variable of interest for mother i. As our main outcome, we focus on a binary indicator of full-time employment measured when children were 2–3 years old (i.e., in the second post-treatment survey) (see Appendix B for details on the definitions of variables). Full-time employment is a dummy variable equal to one if the mother works

[25]To do so, we regress each variable on the treatment indicator, a higher-SES indicator, and their interaction. Column (6) of Table 1 shows p-values of F-tests for joint significance of the coefficients on the treatment indicator and the interaction term.

[26]Hermes et al. (2021) show that also in the first post-treatment survey, attrition is independent of treatment status and is not selective with respect to baseline outcomes.

30 hours or more per week as an employee or when self-employed, and zero otherwise (following the OECD definition of full-time work, e.g., Fluchtmann and Patrini, 2023).[27] As additional outcomes, we consider log monthly net household income, a binary indicator of (part-time or full-time) employment, weekly working hours, and log monthly net earnings of mothers. Moreover, in our last set of analyses we investigate treatment effects on three dimensions of gender inequality in the household (i.e., an indicator of a "male breadwinner household", share of maternal care hours in parental care hours, and share of maternal earnings in parental earnings).

$Treatment_i$ is the treatment indicator, taking a value of one for mothers in the treatment group, and zero for mothers in the control group. $HigherSES_i$ is an indicator for higher-SES mothers, which takes a value of one if the mother has obtained a college entrance qualification ("higher-SES"), and zero otherwise ("lower-SES"). $Treatment_i \times HigherSES_i$ is the interaction of the treatment dummy with the higher-SES dummy. Thus, the causal ITT effect of our intervention on lower-SES mothers is given by β_1, and the effect on higher-SES mothers is given by $\beta_1 + \beta_2$.

As outlined in our pre-analysis plan, we include a vector of control variables, X_i, to increase the precision of our treatment effect estimates. All control variables come from the baseline survey and refer to the mother. The control variables include pre-treatment values of the respective outcome, pre-treatment employment status, age, migration background, log household income, and whether the mother is the primary caregiver for the child.[28] We also include strata controls. In the few cases where control variables have missing values, we impute missings with the sample mean and add imputation dummies.

Finally, ε_i denotes the error term. The inference is based on robust standard errors. The results also hold when using randomization inference or adjusting for multiple hypothesis testing (see Appendix Table E3).

5. Setting the Stage: Treatment Effects on Child Care Participation

To contextualise our main finding, we first present treatment effects on (full-time) child care enrollment. In Section 6, we turn to the effects of our intervention on longer-run effects on mothers' labor supply and gender equality.

Hermes et al. (2021) analyze the short-term effects of the intervention on child care enrollment as measured in the first post-treatment survey nine month after the intervention

[27] Our findings are robust to excluding self-employed mothers, whose working hours may be less accurately measured as they are typically not specified in a contract. Results are available upon request.

[28] All regressions also control for whether the mother or father has participated in the baseline survey.

(i.e., when children are 1–2 years old). They document a substantial treatment effect on child care enrollment for lower-SES families.[29] Reassuringly, we replicate this enrollment effect within the analytical sample of this study (i.e., those who also took part in the second post-treatment survey 18 months after the treatment). Specifically, the intervention increases the probability of lower-SES families enrolling into child care by 14.5 pp ($p = .041$) (see Table 2, Column (1)). As anticipated from the intervention's design, we do not find significant effects for higher-SES families.[30]

Because high-intensity child care use is likely pivotal for taking up full-time employment, we next investigate treatment effects on full-time child care (see Table 2, Columns (2) and (3)). Our intervention increases the likelihood that lower-SES families secure a full-time slot (25 hours or more per week) by 10.4 pp ($p = .082$). This effect is also economically significant, as it translates to 78% of the control group mean. When defining a full-time slot as taking-up at least 30 hours of child care per week, the treatment effect for lower-SES families remains sizable (8.5 pp or 63% of the control group mean), but is not statistically significant ($p = .152$). Column (4) shows that our intervention increases child care usage on average by roughly four hours per week ($p = .044$). Treatment effects for higher-SES families are not statistically significant throughout. In consequence, our intervention strongly reduces the inequality between lower-SES and higher-SES families, closing about half of the SES gaps observed in the control group for the various enrollment outcomes.

[29] They also investigate possible mechanisms of the intervention and document that the intervention significantly increases application knowledge and behavior (e.g., on-site visits to child care centers) for lower-SES parents. In contrast, the intervention does not affect parental preferences for child care.

[30] We do not find evidence for the existence of displacement effects, i.e., the elevated enrollment rates among treated families do not diminish enrollment chances for families in the control group living nearby. For details, see Hermes et al. (2021).

Table 2: Treatment Effects on Child Care Enrollment and Hours in Child Care (Child Age 1–2Y)

	Child Care Enrollment (1)	Full-Time Care (25h or more) (2)	Full-Time Care (30h or more) (3)	Hours in Child Care (4)
Treatment	0.145**	0.104*	0.085	4.070**
	(0.071)	(0.060)	(0.059)	(2.016)
Treatment × Higher-SES	-0.160*	-0.167**	-0.149*	-5.822**
	(0.092)	(0.081)	(0.080)	(2.688)
Higher-SES	0.226***	0.195***	0.123**	7.187***
	(0.070)	(0.062)	(0.061)	(2.039)
Strata Controls	Yes	Yes	Yes	Yes
Sociodemographic Controls	Yes	Yes	Yes	Yes
Treatment Effect-Higher-SES	-0.015	-0.063	-0.063	-1.752
	(0.060)	(0.057)	(0.055)	(1.854)
Control Mean Higher-SES	0.496	0.370	0.296	13.674
Control Mean Lower-SES	0.220	0.134	0.134	5.244
Control Mean SES Gap	0.277	0.236	0.162	8.430
N	460	460	460	460

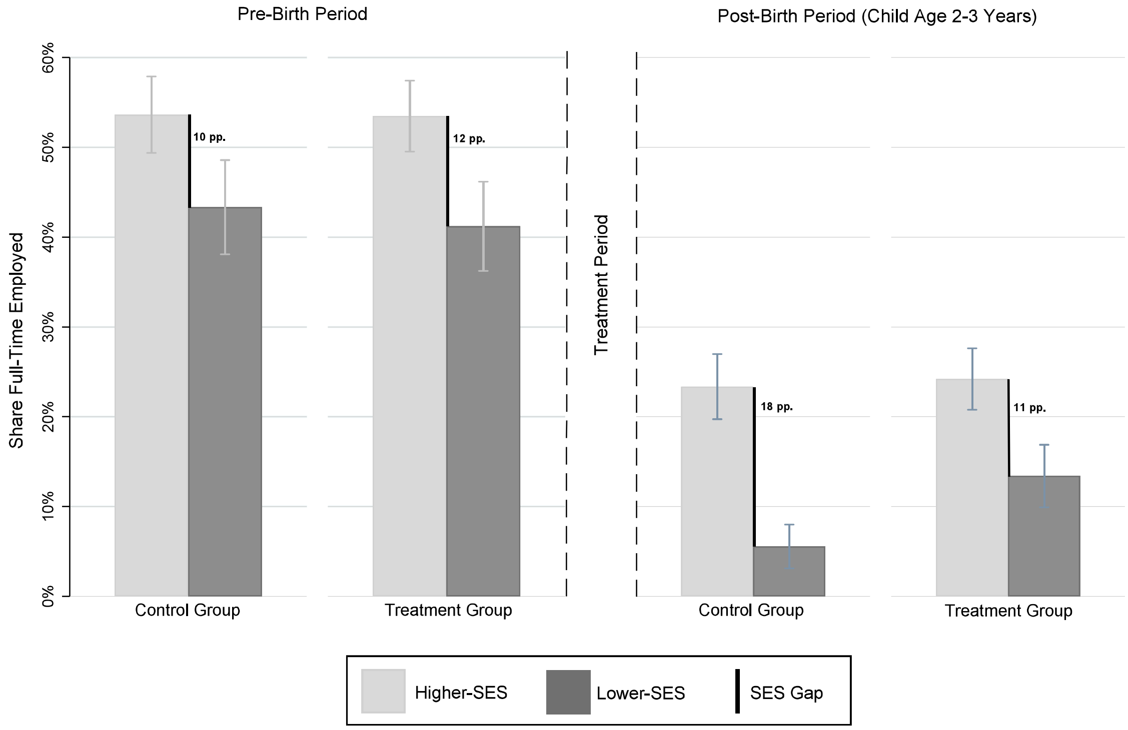
Notes: Table shows intention-to-treat effects on child care enrollment and hours in child care, all models are estimated by OLS. Outcomes are measured nine months after the intervention (i.e., when children are 1–2 years old). 21 mothers did not participate in the first post-treatment survey in which the outcomes were elicited. In Column (1), the outcome takes a value of one if the child is enrolled in child care, zero otherwise. In Column (2) (Column (3)), the outcome takes a value of one if the child visits child care for 25 (30) hours or more per week, zero otherwise. In Column (4), the outcome is weekly hours in child care. *Higher-SES* equals one if the mother has a college entrance qualification, zero otherwise. All models include the pre-treatment enrollment status, strata controls, baseline sociodemographic controls, and survey date fixed effects (see Section 4 and Hermes et al. (2021) for details). Imputation dummies for missing values in control variables are included. *Control Mean Higher-SES* (*Lower-SES*) is the mean of the outcome in the control group in the post-treatment survey for higher-SES (lower-SES) mothers; *Control Group SES Gap* reports the difference between control means of higher- and lower-SES mothers. Robust standard errors in parentheses, * $p<.10$, ** $p<.05$, *** $p<.01$.

6. Treatment Effects on Maternal Labor Market Outcomes

6.1. Full-Time Employment

This section presents treatment effects on full-time employment, our main outcome of interest. We start by comparing the full-time employment rates between treatment and control groups before and after childbirth, separately for higher- and lower-SES mothers, in the raw data. The left part of Figure 2 shows the share of full-time employed mothers in the year before childbirth (i.e., pre-treatment). There are large SES gaps in full-time employment: In the control group, 54% of higher-SES mothers work full-time pre-birth (light grey bars), compared to only 43% of lower-SES mothers (dark grey bars). Thus, we

Figure 2: Maternal Full-time Employment Before vs. After Treatment by SES



Notes: Figure shows the share of full-time employed mothers in the control group and the treatment group before and after childbirth in the raw data. Light grey bars (dark grey bars) indicate the share of higher-SES (lower-SES) mothers in full-time employment. The black lines illustrate the difference in full-time employment between higher- and lower-SES mothers ("SES Gap"). The left part of the figure refers to the year before the child was born. The right part of the figure refers to the post-treatment period, when children are 2–3 years old (18 months after the treatment). Error bars show standard errors.

observe a pre-birth SES gap of 10 pp in full-time employment. In the treatment group, we observe very similar full-time employment rates for lower- and higher-SES mothers and a similar SES gap as in the control group (all treatment-control differences: $p > .750$).

In the right part of Figure 2, we plot full-time employment rates 18 months after the treatment, when children are 2–3 years old. In the control group, the full-time employment rate for higher-SES mothers declines sharply to 23% (a decrease by 57% compared to the pre-birth value). However, the decrease is even stronger for lower-SES mothers, whose full-time employment rates drop to just 5.6% — this corresponds to a decrease by 87%.[31] In consequence, the control-group SES gap in full-time employment increases sharply after childbirth, from 10 to 18 pp, a pattern consistent with previous findings (BPB, 2021).

The treatment strongly mitigates the negative effect of childbirth on full-time employment for lower-SES mothers. The full-time employment rate of lower-SES mothers in the treatment group is 13.4%, about 2.5 times as large as in the control group. Our treatment does not affect full-time employment rates of higher-SES mothers. This finding is fully in line with the result that our treatment increased child care enrollment of lower-SES families, but not of higher-SES families. The treatment thus prevents the childbirth-induced increase in the SES gap in full-time employment observed in the control group, as the SES gap in the treatment group remains at about its pre-birth value (11 pp).

In a second step, we estimate regressions based on Equation (1). As a benchmark, Column (1) of Table 3 shows unconditional treatment effects, corresponding to the results in Figure 2. In Columns (2)–(4), we sequentially add mothers' pre-treatment (i.e., pre-birth) outcome, strata controls, and sociodemographic characteristics. Column (4) shows our pre-specified, preferred model. Here, our treatment increases full-time employment rates of lower-SES mothers by 9.2 pp ($p = .033$), a 160% increase relative to the control group mean.[32] This large treatment effect suggests that a sizable share of lower-SES mothers would actually like to work full-time, but is unable to do so due to barriers in access to universal child care. Reassuringly, our treatment does not affect full-time employment of higher-SES mothers, as treatment effects are virtually zero across specifications (see bottom of Table 3). This result is consistent with the absence of treatment effects on child care enrollment for higher-SES families.

From a policy perspective, it is key to learn which mothers respond to the treatment by taking up full-time work. We thus analyze heterogeneous treatment effects for four different subgroups of mothers. We find that our treatment is more effective for mothers (i) who worked full-time in the year before the child was born, (ii) who have the possibility to return to their pre-birth job, (iii) who have reported in the baseline survey that they planned to return to work, and (iv) who planned to return early after childbirth (see

[31] These values closely mirror the situation of mothers in Germany overall. In the German Microcensus from 2019, the share of full-time-employed mothers with a child between 2 and 3 years is almost identical as in our sample (16.9% vs. 16.3%) (BPB, 2021). In the SOEP data, where we can distinguish mothers by SES background, we observe that 17.1% of higher-SES and 7.6% of lower-SES mothers work full-time when their child is 2–3 years old (own calculation based on SOEP, 2019).

[32] The effect remains large and significant for the subsample also participating in the first post-treatment survey ($n = 460$), for which we estimate treatment effects on enrollment in Section 5 (see Appendix Table E2, Column (3)).

Table 3: Treatment Effects on Maternal Full-Time Employment (Child Age 2–3Y)

	Mother Works Full-Time			
	(1)	(2)	(3)	(4)
Treatment	0.078*	0.084**	0.086**	0.092**
	(0.042)	(0.042)	(0.043)	(0.043)
Treatment × Higher-SES	-0.070	-0.075	-0.082	-0.097
	(0.066)	(0.064)	(0.065)	(0.064)
Higher-SES	0.178***	0.158***	0.144***	0.110**
	(0.044)	(0.042)	(0.043)	(0.044)
Pre-Treatment Outcome	No	Yes	Yes	Yes
Strata Controls	No	No	Yes	Yes
Sociodemographic Controls	No	No	No	Yes
Treatment Effect	0.008	0.009	0.004	-0.005
Higher-SES	(0.050)	(0.048)	(0.048)	(0.047)
Control Mean Higher-SES				0.234
Control Mean Lower-SES				0.056
Control Mean SES Gap				0.178
N	481	481	481	481

Notes: Table shows intention-to-treat effects on full-time employment of mothers, all models are estimated by OLS. Full-time employment is defined as working 30 hours or more per week and is measured when children are 2–3 years old (18 months after the treatment). *Higher-SES* equals one if the mother has a college entrance qualification, zero otherwise. Column (1) shows unconditional treatment effects. In Column (2), we control for an indicator of maternal full-time employment in the year before the child was born. In Column (3), we additionally control for strata variables, and in Column (4) we add controls for sociodemographic characteristics of the mother. All control variables were elicited in the baseline survey. See Section 4 for details on the control variables and variable definitions. Imputation dummies for missing values in control variables are included. *Control Mean Higher-SES* (*Lower-SES*) is the mean of the outcome in the control group in the post-treatment survey for higher-SES (lower-SES) mothers; *Control Group SES Gap* reports the difference between control means of higher- and lower-SES mothers. We additionally report p-values based on randomization inference and correcting for multiple hypothesis testing in Panel A of Table E3. Robust standard errors in parentheses, * $p<.10$, ** $p<.05$, *** $p<.01$.

Appendix Figure D1). These results are consistent with the idea that many mothers would actually like to work full-time but are unable to do so due to the lack of (sufficient) child care.

6.2. Additional Labor Market Outcomes

In this section, we study treatment effects on monthly net household income, employment at the extensive margin, weekly working hours, and monthly net earnings.

First, consistent with the sizeable full-time employment effect, Column (1) of Table 4 shows that the treatment increases monthly net household income by 10.4% for lower-SES

Table 4: Treatment Effects on Other Maternal Labor Market Outcomes (Child Age 2–3Y)

			Working Mothers	
	Log Household Income	Employment	Working Hours	Log Earnings
	(1)	(2)	(3)	(4)
Treatment	0.104*	-0.008	4.903**	0.222*
	(0.056)	(0.068)	(2.072)	(0.129)
Treatment × Higher-SES	-0.142**	0.004	-5.769**	-0.217
	(0.068)	(0.086)	(2.627)	(0.172)
Higher-SES	0.167***	0.003	4.318**	0.172
	(0.058)	(0.064)	(2.021)	(0.136)
Pre-Treatment Outcome	Yes	Yes	Yes	Yes
Strata Controls	Yes	Yes	Yes	Yes
Sociodemographic Controls	Yes	Yes	Yes	Yes
Treatment Effect	-0.038	-0.003	-0.866	0.005
Higher-SES	(0.038)	(0.053)	(1.633)	(0.108)
Control Mean Higher-SES	8.227	0.547	25.314	7.211
Control Mean Lower-SES	7.819	0.444	19.000	6.683
Control Mean SES Gap	0.407	0.103	6.314	0.528
N	459	481	223	223

Notes: Table shows intention-to-treat effects on additional labor market outcomes of mothers, all models are estimated by OLS. Outcomes are measured when children are 2–3 years old (18 months after the treatment). In Column (1), *Log Household Income* is missing for $n = 22$ cases in which household income was not reported. In Column (2), *Employment* is a dummy variable equal to one if the mother works part-time or full-time, and zero otherwise. In Columns (3) and (4), *Working Hours* and *Log Earnings* are estimated in the subsample of working mothers ($n = 244$); $n = 21$ mothers with missing information on working hours and earnings are excluded in these regressions. *Higher-SES* is a dummy variable equal to one if the mother has a college entrance qualification, zero otherwise. All models include controls for the pre-treatment outcome, which refer to the year before the child was born (*log household income* in Column (1) refers to the time of the baseline survey); in Column (3), we control for pre-treatment maternal full-time employment because we did not elicit working hours at baseline. All models further include strata variables and baseline sociodemographic controls (see Section 4 for details). Imputation dummies for missing values in control variables are included. *Control Mean Higher-SES* (*Lower-SES*) is the mean of the outcome in the control group in the post-treatment survey for higher-SES (lower-SES) mothers; *Control Group SES Gap* reports the difference between control means of higher- and lower-SES mothers. We additionally report p-values based on randomization inference and correcting for multiple hypothesis testing in Table E3. Robust standard errors in parentheses, * $p<.10$, ** $p<.05$, *** $p<.01$.

families ($p = .065$). This suggests that the increase in maternal full-time employment did not lead to a compensating decrease in the labor supply (and thus earnings) of fathers.[33]

Next, we investigate whether the full-time effect reported in the previous section is driven by increased labor market participation (extensive margin) or by increased working hours (intensive margin). We find no treatment effect on whether mothers work at all (Column (2) of Table 4). Instead, lower-SES mothers in the treatment group work more hours than those in the control group: Conditional on being employed, the treatment

[33] Given that child care is free of charge for all children aged two years and older in the federal state we study, household income is unaffected by child care expenses.

increases weekly working hours by 4.9 hours ($p = .019$, Column (3)). Thus, the treatment effect on full-time employment is clearly driven by changes at the intensive margin.

One possible reason for the lack of treatment effects at the extensive margin is that mothers without child care may well manage to work part-time (possibly by relying on care provided by other family members for a few hours a day), but are only able to work full-time once they obtain a child care slot. In fact, several pieces of evidence are in line with the idea that full-time employment is more strongly affected by the availability of child care than part-time employment. First, the gender gap in full-time employment is much larger than in part-time employment (e.g., when considering parents with two-year-old children in Germany, the gender gap is 31 pp for overall employment, but as large as 66 pp for full-time employment, see BPB, 2021). Second, about half of the mothers in our sample who work part-time do not have their child enrolled in a child care center, while this is true for only about a quarter of mothers who work full-time. Third, in the control group, only 11% of lower-SES mothers who reported in the baseline survey that they want to work full-time were able to fulfill these aspirations, compared to 47% of lower-SES mothers in the treatment group. In contrast, for part-time employment, the proportion of lower-SES mothers realizing their work plans is more similar in the control group (52%) and treatment group (40%). Another potential reason for observing effects only at the intensive employment margin is that it is easier for mothers to increase their working hours in a given job upon gaining access to child care than to find a new job. Consistently, we observe that our treatment is more effective for mothers who indicate having the option to return to their pre-birth job (see Appendix Figure D1).

Finally, Column (4) of Table 4 shows that monthly net earnings of working lower-SES mothers increased by 22% ($p = 0.087$) due to the treatment. The effect size is large, given that the wage returns to a full year of schooling in Germany typically amount to 8-10% (e.g., Hanushek et al., 2015). The effect is driven by maternal labor supply, as we do not find treatment effects on hourly wages of mothers (results available upon request).

Note that our treatment effect estimates in Columns (3) and (4) of Table 4 would be biased if the selection of mothers into employment would differ between treatment and control groups. We consider such differential selection unlikely since employment status is unaffected by the treatment (see above). Moreover, results on maternal working hours and earnings are similar in a Heckman selection model (Heckman, 1979), which uses either the number of children in the household or the existence of another child (born after baseline) as exclusion restriction (see Appendix Table D1).[34]

6.3. Robustness

In Section 3.4, we already discussed that our results are robust to accounting for sample attrition. Below, we provide additional robustness tests.

Selection into the Sample. Utilizing birth registry data for the entire cohort enables us to investigate the degree of generalizability of our results. To do so, we correct our estimates for selection into the sample by applying propensity score weights. These weights are calculated by regressing a binary variable indicating participation in our baseline survey on all variables available in the birth registry data using a probit model. The re-weighted treatment effect estimates closely align with the unweighted estimates (see Appendix Table E1).

Multiple Hypothesis Testing and Randomization Inference. We adjust our treatment effects for multiple hypothesis testing (MHT) and conduct randomization inference (see Young, 2019). Results are presented in Appendix Table E3. All treatment effects on full-time employment and working hours for lower-SES families remain significant at the 10%-level or better when applying three different MHT corrections (suggested by List et al., 2019; Westfall and Young, 1993; Romano and Wolf, 2005, 2016) and randomization inference, respectively. The treatment effects on household income and maternal earnings remain significant at the 10%-level when applying randomization inference, and are significant at the 15%-level or better for the MHT corrections.

Alternative Definitions of Lower-SES Background. Our preferred definition of maternal SES background is based on maternal education, following previous literature investigating maternal labor supply (e.g., Dujardin et al., 2018; Kuziemko et al., 2018; Müller and Wrohlich, 2020). In particular, we define a mother as having a "lower-SES" background when she does not have a college entrance qualification (henceforth "SES-1", 39% of the sample). We test three alternative definitions of SES: (i) neither of the two parents has a college entrance qualification ("SES-2", 30% of the sample), (ii) either "SES-1" or equivalent household income at baseline is below the poverty line ("SES-3", 51% of the sample, following Falk et al., 2021), and (iii) either "SES-3" or single mother status ("SES-4",

[34]Our results for household income, working hours, and earnings are qualitatively similar when winsorizing the outcome variables to account for potential outliers.

52% of the sample, following Kosse et al., 2020). As shown in Appendix Table E4, treatment effects are robust to including education of the father, household income, and single mother status in our SES definition.

6.4. Mechanisms

This section aims to shed light on the mechanisms driving the treatment effect on maternal full-time employment in the lower-SES sample. As our intervention was designed to increase early child care use, our first potential mediator is hours in child care centers.[35] Since the treatment may also affect how families combine different forms of non-maternal child care, we use care hours provided by fathers or by other caretakers (i.e., hours with grandparents, older siblings, other relatives, friends, nannies, and childminders) as additional potential mediators.[36] Such additional non-maternal care arrangements may matter because opening hours of child care centers are not necessarily covering the full working day. This issue seems to be particularly relevant for lower-SES mothers, as they are more likely to work non-standard schedules, e.g., in shifts or at night (Han et al., 2020).[37]

Our analysis of mechanisms proceeds as follows: First, we analyze treatment effects on the potential mediators, measured in the first post-treatment survey (i.e., nine months before we measure full-time employment, when the child is 1–2 years old). The child care arrangements at this time provide the conditions under which families can plan and organize their future work arrangements. Second, we follow the approach developed by Heckman et al. (2013) to decompose the overall treatment effect into shares attributed to the different mediators.[38]

Treatment Effects on Non-Maternal Care Hours. In a first step, we investigate treatment effects on our potential mediators. Column (1) of Appendix Table F1 shows that the

[35] Above, for exposition, the term "child care" typically referred to center-based child care. However, in this section, we explicitly use the term "center-based child care" to distinguish child care in child care centers from other forms of child care.

[36] Note that these variables always refer to the number of hours a child spends in the respective care arrangement without the involvement of other caretakers. That is, child care by fathers refers to the number of hours the father takes care of the child alone.

[37] Another potential channel through which our treatment may affect full-time employment of mothers is fertility (see, e.g., Doepke and Kindermann, 2019). However, we find no treatment effects of fertility rates.

[38] In the mediation analysis, we add fixed effects for the date of conducting the first post-treatment survey as additional controls to account for the importance of timing in the child care slot allocation process (e.g., child care centers may allocate slots based on the application date or allocate available slots after certain deadlines, see Hermes et al., 2021). As each of our surveys was in the field for several weeks, treatment effects on child care enrollment and on other non-maternal care arrangements may be affected by timing effects.

number of weekly hours a child spends in child care increases in treated families as compared to control families by 4 hours ($p = .049$). Intriguingly, we find that fathers get more involved in child care in treated lower-SES families, as care hours by fathers increase by 2.6 hours per week ($p = .074$) (Column (2)). One potential explanation for this effect is that child care centers may not cover the mothers' entire working day, requiring fathers to cover the remaining hours to enable maternal full-time employment.

We also find that treated lower-SES families have to rely less on alternative care arrangements. There is a negative treatment effect on hours provided by other caretakers in lower-SES families of 2.3 hours per week, albeit being only marginally significant ($p = 0.099$) (Column (3) of Appendix Table F1). Auxiliary analysis shows that this effect is driven by fewer hours provided by family members other than the father (i.e., grandparents, other relatives, or older siblings). Thus, the treatment reduced the dependence on alternative care arrangements beyond center-based child care and the core family.[39] In sum, the treatment increased the number of care hours provided by child care centers and fathers, thus relieving mothers from some of their child care duties and potentially allowing them to work full-time.

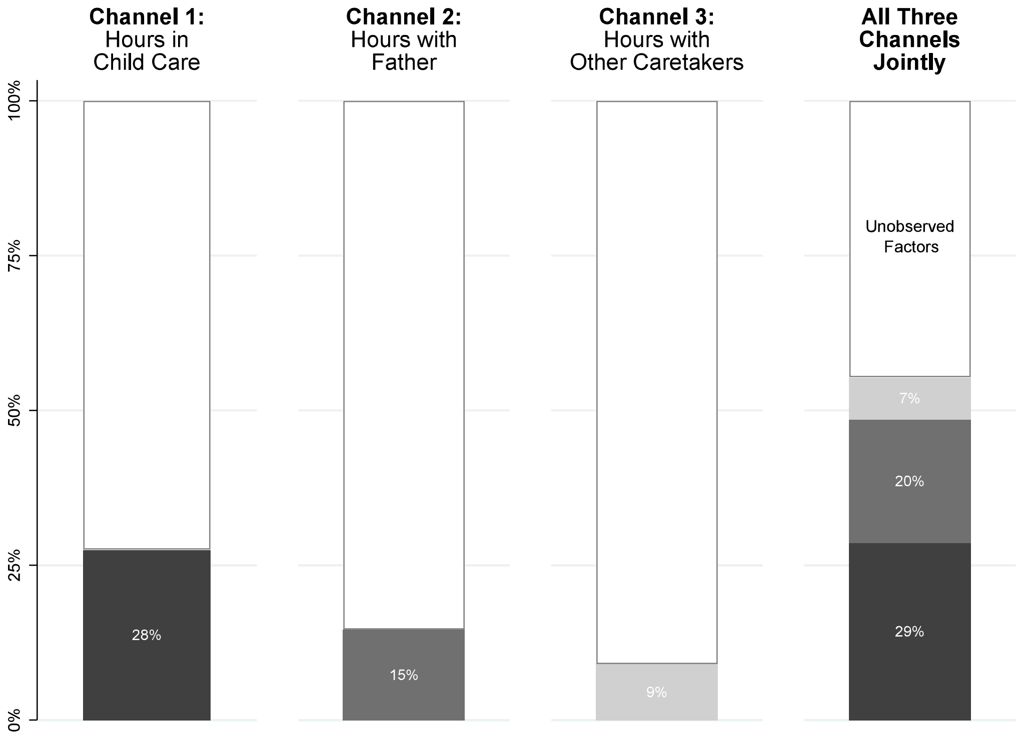
Mediation Analysis. We then conduct a mediation analysis to investigate the share of the treatment effect on full-time employment that our mediators can explain. We follow the approach by Heckman et al. (2013) and Heckman and Pinto (2015) (applied, e.g., by Oreopoulos et al., 2017; Kosse et al., 2020; Resnjanskij et al., 2023) and decompose the treatment effect into the shares explained by our observed mediator variables and a remaining share explained by unobserved mediators. Here we report the main results of the mediation analysis, while the methodological details are relegated to Appendix F.[40]

Since the treatment significantly affects all potential mediators discussed above, we add them to our baseline regression shown in Equation (1). More precisely, we include (i) hours in center-based child care, (ii) hours with the father, and (iii) hours with other caretakers. Appendix Table F2 shows how the inclusion of the mediator variables affects treatment effects on maternal full-time employment. We combine these results with the

[39]In line with our previous results, we find no treatment effects on non-maternal care hours for higher-SES mothers.

[40]An alternative to the mediation analysis would be an instrumental variable (IV) estimation to assess the role of child care enrollment for maternal full-time employment. The IV estimates suggest a positive effect of child care enrollment on full-time employment rates of lower-SES mothers. However, we refrain from interpreting this result because of the weak first stage ($F = 3.24$) due to the small sample size.

Figure 3: Treatment Effects on Maternal Full-Time Employment (Child Age 2–3Y) with Mediators



Notes: Figure shows the share of the treatment effect on full-time employment for lower-SES mothers measured when children are 2–3 years old (18 months after the treatment) that can be attributed to the respective mediator measured when children are 1–2 years old (nine months after the treatment). The first three bars show the contribution of a single mediator, while the last bar shows the contribution of all three mediators when they are jointly included. The mediation analysis is based on $n = 436$ observations for which information on all mediators is available. Detailed results are reported in the Appendix Tables F1 and F2.

estimated treatment effect on the respective mediator, provided in Appendix Table F1, to assess the relative contribution of each mediator in explaining the full-time effect.

Figure 3 illustrates the results of the mediation analysis. In the first three bars, we consider each mediator separately. We find that increased hours in center-based child care explain 28% of the treatment effect on full-time employment for lower-SES mothers; increased hours with the father explain 15%, and decreased hours with other caretakers explain 9%. Including all three mediators jointly in the last bar of the figure, we can explain as much as 56% of the treatment effect. The largest shares can be attributed to increases in hours provided by center-based child care (29%) and fathers (20%), while 7% can be attributed to the decreased dependence on other care arrangements.

In sum, the mechanism analysis provides the intuitive result that our treatment effect mainly materializes through an increase in non-maternal care hours, which allows mothers to work full-time by relieving them from some of the child care obligations. This finding corroborates the non-experimental result of Andresen and Nix (2022a) that providing alternatives to maternal care hours is an effective strategy to reduce child penalties for (lower-SES) mothers.

7. Treatment Effects on Gender Equality

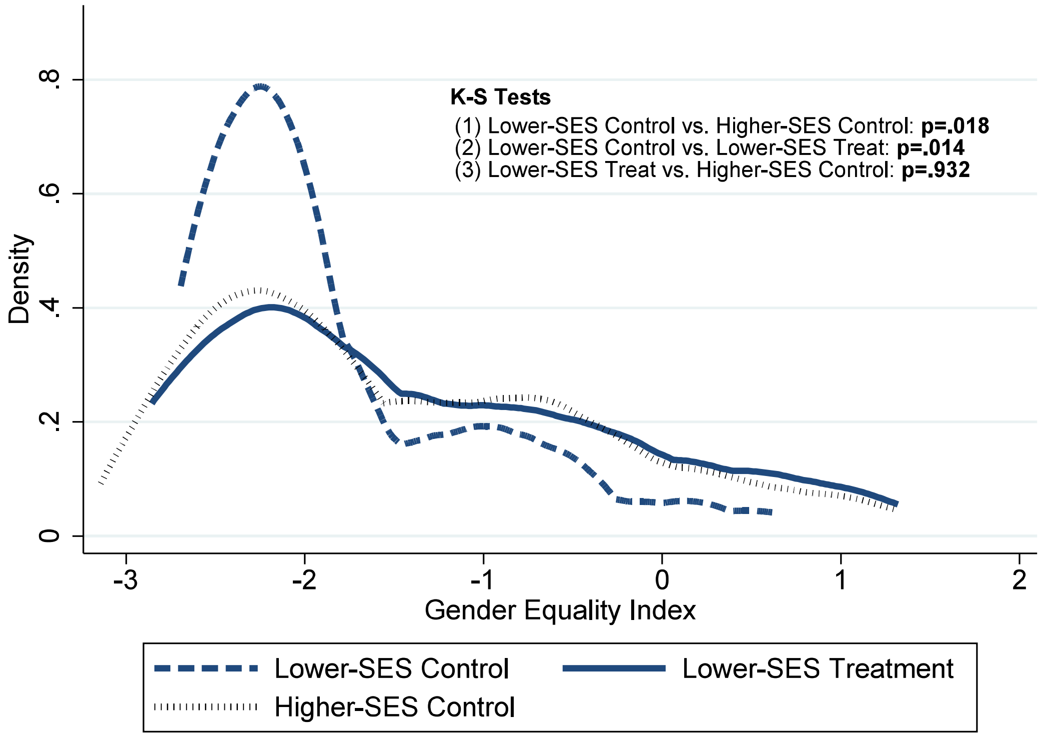
Our mechanism analysis suggests that enabling families to access universal child care encourages fathers to assume greater child care responsibilities. This finding aligns with the broader objective of enhancing gender equality, a key aim of child care and family policies in developed countries (OECD, 2011, 2012, 2017b; Lundberg, 2023). We now proceed to directly examine treatment effects on (intra-household) gender equality.

Leveraging information on both mothers and fathers in our survey data, we compute an index of intra-household gender equality that consists of three elements: (i) "male breadwinner" households (i.e., the father works full-time, while the mother is employed part-time or not employed at all); (ii) share of maternal care hours in total parental care hours; and (iii) share of maternal earnings in total household earnings (see Appendix B for details).[41] Importantly, these three outcomes measure different dimensions of intra-household gender equality: Correlations between these outcomes in the control group are small to moderate ($|\rho| = .15 - .52$), especially in the lower-SES sample ($|\rho| = .05 - .34$). To reduce measurement error and alleviate multiple testing concerns, we combine these three outcomes into a gender equality index (Kling et al., 2007), standardized such that a value of zero indicates a household with full gender equality (see notes to Table 5).

Figure 4 summarizes our key findings regarding treatment effects on gender equality: First, in the control group, lower-SES households exhibit substantially lower levels of gender equality than higher-SES households (K-S test of the equality of distributions, $p =$ .018). Second, the treatment has a profound impact on lower-SES households, significantly shifting their distribution towards higher levels of gender equality (K-S test, $p = .014$). Third, the distribution of gender equality among the treated lower-SES households closely resembles that of the higher-SES households in the control group (K-S test, $p = .932$).

[41] Appendix B shows that the gender gap in earnings in our data aligns very well with the gap found in nationally representative data.

Figure 4: Treatment Effect on the Distribution of Gender Equality for Lower-SES Mothers



Notes: Figure shows the unconditional treatment effect on the distribution of the gender equality index for lower-SES mothers. The dark blue dashed (solid) line shows the distribution for lower-SES mothers in the control (treatment) group. For comparison, the black vertically dashed line depicts the distribution of the gender equality index for higher-SES mothers in the control group. The probability density functions are computed with an Epanechnikov kernel with bandwidth h derived from the Silverman rule (Silverman, 1986) with $h = 0.9An^{-\frac{1}{5}}$, where n is the number of observations and $A = min(standard\,deviation, interquartile\,range/1.349)$. K-S Tests: p-values for a Kolmogorov-Smirnov test of the equality of distributions of the gender equality index of (1) lower-SES mothers and higher-SES mothers in the control group, (2) lower-SES mothers in the treatment and control groups and (3) lower-SES mothers in the treatment group and higher-SES mothers in the control group.

This suggests that the treatment effectively closes the existing SES gap in overall gender equality.

Similar to our findings for other outcomes, there is no treatment effect on the gender equality index for higher-SES households. In fact, the distributions of gender equality in the treatment and control groups for higher-SES households are almost identical (see Appendix Figure D2).

While Figure 4 provides an unconditional comparison of gender equality between treatment and control groups, we also estimate treatment effects based on the empirical model

Table 5: Treatment Effects on Gender Equality in the Household

	Gender Equality Index (1)	Male Breadwinner Household (2)	Share Maternal Care Hours (3)	Share Maternal Earnings (4)
Treatment	0.401***	-0.159**	-0.111**	0.070*
	(0.147)	(0.068)	(0.052)	(0.040)
Treatment × Higher-SES	-0.328*	0.142	0.104	-0.035
	(0.187)	(0.088)	(0.066)	(0.050)
Higher-SES	0.197	-0.128*	-0.062	0.027
	(0.130)	(0.066)	(0.044)	(0.034)
Pre-Treatment Outcome	Yes	Yes	Yes	Yes
Strata Controls	Yes	Yes	Yes	Yes
Sociodemographic Controls	Yes	Yes	Yes	Yes
Treatment Effect	0.073	-0.016	-0.007	0.035
Higher-SES	(0.118)	(0.055)	(0.042)	(0.028)
Increase in Gender Equality (in %)		19.8	30.5	32.0
Control Mean Higher-SES	-1.435	0.646	0.773	0.358
Control Mean Lower-SES	-1.834	0.803	0.864	0.281
Control Mean SES Gap	0.399	-0.157	-0.091	0.077
N	441	441	405	206

Notes: Table shows intention-to-treat effects on the intra-household gender equality index, as well as its three constituent dimensions, estimated by OLS. All models exclude single mothers, as we cannot compute intra-household gaps for them ($n = 40$ mothers in Column (1), $n = 40$ mothers in Column (2), $n = 31$ mothers in Column (3), and $n = 15$ mothers in Column (4)). In Column (1), *Gender Equality Index* is an index that combines all three dimensions of gender equality: *Male Breadwinner Household*, *Share Maternal Care Hours* and *Share Maternal Earnings* (all standardized, assigning equal weights to each variable). The index is standardized to mean = -1.58 and SD = 1 in the control group (Kling et al., 2007), such that a value of zero indicates a household with gender equality. In Column (2), *Male Breadwinner Household* is a dummy variable taking a value of one if the father is employed full-time and the mother is employed part-time or not employed, and zero otherwise; information is obtained from the second post-treatment survey. In Column (3), *Share Maternal Care Hours* is calculated by dividing maternal care hours by the sum of care hours provided by the mother and father; information is obtained from the first post-treatment survey (see Section 6.4). The number of observations decreases in this specification, because we exclude mothers who did not participate in the first post-treatment survey ($n = 21$) and those who did not answer the question on care hours ($n = 24$). In Column (4), *Share Maternal Earnings* is calculated by dividing earnings of mothers by total household earnings (following Bertrand et al., 2015), excluding mothers who are not working; information is obtained from the second post-treatment survey. *Higher-SES* equals one if the mother has a college entrance qualification, zero otherwise. All models include controls for pre-treatment outcomes; we use pre-birth earnings of the mother in Column (4) because we do not have information on pre-birth earnings of the father. Additionally, we include controls for strata variables and baseline sociodemographic controls (see Section 4 for details). Imputation dummies for missing values in control variables are included. In Column (2), *Increase in Gender Equality (in %)*, shown in the bottom of the table, is calculated by dividing the treatment effect for lower-SES mothers by the mean of the outcome in the lower-SES control group. In Column (3) and (4), *Increase in Gender Equality (in %)* is calculated by dividing the treatment effect for lower-SES mothers by the gender gap in the lower-SES control group; the gender gap is calculated as the deviation of the lower-SES control group mean from an equal division of care hours and earnings, respectively. *Control Mean Higher-SES* (*Lower-SES*) is the mean of the outcome in the control group in the post-treatment survey for higher-SES (lower-SES) mothers; *Control Group SES Gap* reports the difference between control means of higher- and lower-SES mothers. We additionally report p-values based on randomization inference and correcting for multiple hypothesis testing for Columns (2)–(4) in Panel C of Table E3. Robust standard errors in parentheses, * $p< .10$, ** $p< .05$, *** $p< .01$.

in Equation (1). Table 5 confirms the large, equity-enhancing effects of our treatment for lower-SES families. In Column (1), the gender equality index increases by 40% of a standard deviation due to our treatment ($p = .007$). Remarkably, this effect precisely matches the size of the SES gap in the control group, reported at the bottom of the table. The treatment effect on higher-SES families is small and insignificant. Thus, the treatment fully closes the SES gap in gender equality, as visually depicted in Figure 4.

In Columns (2)–(4) of Table 5, we show that each dimension of the gender equality index is significantly improved by the treatment. We find that treated lower-SES families (i) are 15.9 pp less likely to be a "male breadwinner" household ($p = .020$), (ii) experience a 11.1 pp reduction in maternal care hours as a share of total parental care hours ($p = .034$), and (iii) exhibit an increase in maternal earnings as a share of household earnings by 7.0 pp ($p = .087$).[42] These treatment effect sizes correspond to improvements in gender equality by 19.8%, 30.5%, and 32.0%, respectively (see bottom of Table 5).[43]

Our results indicate that access to universal child care promotes a more gender-equal distribution of paid *and* unpaid work within lower-SES households. The more equal distribution of full-time work and earnings could have important long-run implications, potentially leading to a convergence in labor market opportunities and pension entitlements between mothers and fathers, a more balanced bargaining power, and perhaps even to more progressive gender attitudes among their children (see, e.g., Farré et al., 2021).[44]

8. Conclusion

We present the first randomized controlled trial that investigates the effect of enabling access to universal early child care for lower-SES families on maternal labor market outcomes and intra-household gender equality. Our intervention offered families with children below one year information and customized assistance, which substantially increased (full-time) child care enrollment for lower-SES families. We find that the labor market success

[42] When correcting for multiple hypothesis testing along the three dimensions of gender equality, results remain significant at the 10%-level. However, for the share of maternal earnings in household earnings, treatment effects are significant only for one out of three correction methods (see Appendix Table E3.)

[43] For instance, in the lower-SES control group, mothers take over 86.4% of child care work, i.e., 36.4 pp more than suggested by an equal division of caregiving duties. The treatment reduces this gap by 11.1 pp or, in relative terms, 30.5% (see Column (3) of Table 5).

[44] A potential concern could be that mothers are not satisfied with their increased labor market activity. However, this concern is not supported in our data. In a hypothetical scenario, we ask mothers independently of their actual working status how taking up full-time work would affect their well-being. Treated mothers regard full-time employment as more beneficial for their well-being than control-group mothers ($p = .035$).

of lower-SES mothers strongly benefitted from the treatment in various dimensions: Compared to the control group, lower-SES mothers in the treatment group are more likely to work full-time (+9.2 pp), have a higher household income (+10%), work more hours (+5 hours per week), and have higher earnings (+22%). As a consequence, the treatment substantially improves intra-household gender equality: An overall index of gender equality consisting of working hours, care hours, and earnings increases by as much as 40% of a standard deviation. Intriguingly, we find improvements in gender equality in each dimension of the index.

From a policy perspective, our findings suggest that providing easier access to universal early child care is an effective policy tool for increasing full-time employment rates of lower-SES mothers — a group that is particularly disadvantaged in the labor market. Activating the working potential of mothers seems more important than ever in light of the immense labor shortage in aging societies like Germany (e.g., more than 50% of German companies are currently affected by labor shortages; see Peichl et al., 2022). Moreover, by decreasing maternal child care obligations, an improved access to universal early child care also alleviates mothers' time constraints, allowing them to make labor market choices that better match their preferences. Our findings on intra-household gender equality also indicate that expanding universal early child care access can accelerate the transition to a more gender-equal society. And while such expansion of child care comes at considerable public expenses, our results suggest that fiscal considerations should not discourage policy-makers from improving child care access. In fact, a back-of-the-envelope calculation shows that the treatment-induced increase in income tax revenue would be sufficient to finance the additional costs of providing more child care hours (see Appendix G for details).

Our results also have implications for the effective design of social programs more generally. Administrative burdens are a major barrier to accessing social benefits and (formally non-selective) educational programs, especially for lower-SES individuals (see, e.g., Currie, 2006; Walters, 2018; Ko and Moffitt, 2022). We demonstrate that removing such barriers (i) promotes child care take up by reducing inefficiencies in the existing system of allocating child care slots (discussed in detail in Hermes et al., 2021) and (ii) yields large benefits by improving mothers' labor market opportunities. The evidence presented in this paper not only calls for additional efforts by policy-makers to increase the number of child care slots. It also emphasizes the need of structural reforms to reduce barriers in the application process (e.g., reducing paperwork or centralizing the admission system) to fully realize the societal benefits of universal early child care programs.

References

Ajayi, K. F., A. Dao, and E. Koussoubé (2022). The Effects of Childcare on Women and Children: Evidence from a Randomized Evaluation in Burkina Faso. CGD Working Paper 30653.

Andresen, M. and T. Havnes (2019). Child Care, Parental Labor Supply and Tax Revenue. *Labour Economics 61*, 101762.

Andresen, M. and E. Nix (2022a). Can the Child Penalty be Reduced?. Evaluating Multiple Policy Interventions. Discussion Papers 983, Statistics Norway, Research Department.

Andresen, M. and E. Nix (2022b). What Causes the Child Penalty? Evidence from Adopting and Same Sex Couples. *Journal of Labor Economics 40*(4), 971–1004.

Athey, S. and G. W. Imbens (2017). The Econometrics of Randomized Experiments. In A. V. Banerjee and E. Duflo (Eds.), *Handbook of Economic Field Experiments*, Volume 1, pp. 73–140. North Holland, Amsterdam.

Attanasio, O., R. Paes de Barros, P. Carneiro, D. K. Evans, L. Lima, P. Olinto, and N. Schady (2022). Public childcare, labor market outcomes of caregivers, and child development: Experimental evidence from brazil. NBER Working Paper 30653, NBER.

Baker, M., J. Gruber, and K. Milligan (2008). Universal Child Care, Maternal Labour Supply, and Family Well-Being. *Journal of Political Economy 116*(4), 709–745.

Bauernschuster, S. and M. Schlotter (2015). Public Child Care and Mothers' Labor Supply - Evidence from Two Quasi-Experiments. *Journal of Public Economics 123*, 1–16.

Berniell, I., L. Berniell, D. de la Mata, M. Edo, and M. Marchionni (2021). Gender Gaps in Labor Informality: The Motherhood Effect. *Journal of Development Economics 150*, 102599.

Bertrand, M., E. Kamenica, and J. Pan (2015). Gender Identity and Relative Income within Households. *The Quarterly Journal of Economics 130*(2), 571–614.

Bjorvatn, K., D. Ferris, S. Gulesci, A. Nasgowitz, and V. Somville (2022). Childcare, labor supply, and business development: Experimental evidence from Uganda. CEPR Press Discussion Paper No. 17243, Centre for Economic Policy Research.

BMFSFJ (2012). Ausgeübte Erwerbstätigkeit von Müttern. Technical report, Bundesministerium für Familien, Senioren, Frauen und Jugend (German Ministry of Family Affairs).

BMFSFJ (2020). *(Existenzsichernde) Erwerbstätigkeit von Müttern.* Bundesministerium für Familie, Senioren, Frauen und Jugend.

Borowsky, J., J. H. Brown, E. E. Davis, C. Gibbs, C. M. Herbst, A. Sojourner, E. Tekin, and M. J. Wiswall (2022). An equilibrium model of the impact of increased public investment in early childhood education. NBER Working Paper 30140, National Bureau of Economic Research.

BPB (2021). Erwerbstätigkeit von Eltern nach Alter des jüngsten Kindes. https://www.bpb.de/nachschlagen/zahlen-und-fakten/soziale-situation-in-deutschland/61606/erwerbstaetigkeit-nach-alter-des-juengsten-kindes, Bundeszentrale für politische Bildung.

Carta, F. and L. Rizzica (2018). Early Kindergarten, Maternal Labor Supply and Children's Outcomes: Evidence from Italy. *Journal of Public Economics 158*, 79–102.

Cascio, E. (2009). Maternal Labor Supply and the Introduction of Kindergartens into American Public Schools. *Journal of Human Resources 44*(1), 140–170.

Cascio, E. (2021). Early Childhood Education in the United States: What, When, Where, Who, How, and Why. NBER Working Paper 28722, National Bureau of Economic Research.

Clark, S., C. W. Kabiru, S. Laszlo, and S. Muthur (2019). The impact of childcare on poor urban women's economic empowerment in africa. *Demography 56*, 1247–1272.

Clarke, D., J. P. Romano, and M. Wolf (2020). The Romano-Wolf Multiple-Hypothesis Correction in Stata. *The Stata Journal 20*(4), 812–843.

Cornelissen, T., C. Dustmann, A. Raute, and U. Schönberg (2018). Who Benefits from Universal Child Care? Estimating Marginal Returns to Early Child Care Attendance. *Journal of Political Economy 126*(6), 2356–2409.

Currie, J. (2006). The Take Up of Social Benefits. In A. Auerbach, D. Card, and J. Quigley (Eds.), *Poverty, the Distribution of Income, and Public Policy*, pp. 80–148. Russell Sage, New York.

Destatis (2021). Bildungsfinanzbericht 2021. German Federal Statistical Office, Wiesbaden.

Destatis (2021). Parental allowance 2020: proportion of fathers up to just under 25%. https://www.destatis.de/en/press/2021/03/pe21_146_22922.html. German Federal Statistical Office, Wiesbaden.

Destatis (2022). Personen in Elternzeit. https://www.destatis.de/de/themen/arbeit/arbeitsmarkt/qualitaet-arbeit/dimension-3/elternzeit.html. German Federal Statistical Office, Wiesbaden.

Destatis (2023). Closing the gap? Erwerbstätigkeit und Arbeitszeit von Müttern und Vätern nach 15 Jahren Elterngeld. https://www.destatis.de/de/methoden/wista-wirtschaft-und-statistik/2023/04/erwerbstaetigkeit-arbeitszeit-042023.html. German Federal Statistical Office, Wiesbaden.

Doepke, M. and F. Kindermann (2019). Bargaining over Babies: Theory, Evidence, and Policy Implications. *American Economic Review 109*(9), 3264–3306.

Dujardin, C., M. Fonder, and B. Lejeune (2018). Does Formal Child Care Availability for 0 - 3 Year Olds Boost Mothers' Employment Rate? Panel Data Based Evidence from Belgium. *Annals of Economics and Statistics 129*, 103–126.

Education Report (2020). Autorengruppe Bildungsberichterstattung: Bildung in Deutschland 2020. Ein indikatorengestützter Bericht mit einer Analyse zur Bildung in einer digitalisierten Welt.

Eika, L., M. Mogstad, and B. Zafar (2019). Educational Assortative Mating and Household Income Inequality. *Journal of Political Economy 127*(6), 2795–2835.

Falk, A., F. Kosse, P. Pinger, H. Schildberg-Hoerisch, and T. Deckers (2021). Socio-Economic Status and Inequalities in Children's IQ and Economic Preferences. *Journal of Political Economy 129*(9), 2504–2545.

Farré, L., C. Felfe, L. González, and P. Schneider (2021). Changing Gender Norms across Generations: Evidence from a Paternity Leave Reform. BSE Working Paper 1310, Barcelona School of Economics.

Felfe, C. and R. Lalive (2018). Does Early Child Care Affect Children's Development? *Journal of Public Economics 159*, 33–53.

Felfe, C., N. Nollenberger, and N. Rodríguez-Planas (2015). Can't buy mummy's lov? Universal childcare and children's long-term cognitive development. *Journal of Population Economics 28*(2), 393–422.

Fitzpatrick, M. D. (2010). Preschoolers Enrolled and Mothers at Work? The Effects of Universal Prekindergarten. *Journal of Labor Economics 28*(1), 51–85.

Flisi, S., Z. Blasko, and E. Stepanova (2022). Indicators for early childhood education and care – Reconsidering some aspects of the Barcelona target for younger children. Luxembourg, Publications Office of the European Union.

Flood, S., J. F. McMurry, A. Sojourner, and M. J. Wiswall (2022). Inequality in Early Care Experienced by U.S. Children. *Journal of Economic Perspectives 36*(2), 199–222.

Fluchtmann, J. and V. Patrini (2023). Women at work in OECD countries. In OECD (Ed.), *Joining Forces for Gender Equality: What is Holding us Back?*, pp. 145–155. OECD Publishing, Paris.

Geis-Thöne, W. (2021). Mütter haben unterschiedliche Erwerbswünsche und erwerbsbezogene Normen. IW-Report 28/2021, IW, Köln.

Goebel, J., M. M. Grabka, S. Liebig, M. Kroh, D. Richter, C. Schroeder, and J. Schupp (2019). The German Socio-Economic Panel (SOEP). *Journal of Economics and Statistics 239*(2), 345–360.

Goldin, C. (2021). *Career and Family: Women's Century-Long Journey toward Equity.* New Jersey: Princeton University Press.

Goldin, C. (2023). Why Women Won. NBER Working Paper 31762, National Bureau of Economic Research.

Goux, D. and E. Maurin (2010). Public School Availability for Two-Year Olds and Mothers' Labour Supply. *Labour Economics 17*(6), 951–962.

Han, W.-J., P. Gracia, and J. Li (2020). Parental Work Schedules and Hours in 29 European Countries, 2005-2015: A Welfare State Comparison. *SocArXiv*.

Hanushek, E. A., G. Schwerdt, S. Wiederhold, and L. Woessmann (2015). Returns to skills around the world: Evidence from piaac. *European Economic Review 73*, 103–130.

Harnisch, M., K.-U. Müller, and M. Neumann (2018). Teilzeitbeschäftigte würden gerne mehr Stunden arbeiten, Vollzeitbeschäftigte lieber reduzieren. *DIW Wochenbericht 38*, 837–846.

Havnes, T. and M. Mogstad (2011). No child left behind: Subsidized child care and children's long-run outcomes. *American Economic Journal: Economic Policy 3*(2), 97–129.

Heckman, J. (1979). Sample Selection Bias as a Specification Error. *Econometrica 47*(1), 153–161.

Heckman, J. and R. Landersø (2022). Lessons for Americans from Denmark about Inequality and Social Mobility. *Labor Economics 77*, 101999.

Heckman, J., R. Pinto, and P. Savelyev (2013). Understanding the Mechanisms Through Which an Influential Early Childhood Program Boosted Adult Outcomes. *American Economic Review 103*(6), 2052–2086.

Heckman, J. J. and R. Pinto (2015). Econometric Mediation Analyses: Identifying the Sources of Treatment Effects from Experimentally Estimated Production Technologies with Unmeasured and Mismeasured Inputs. *Econometrics Review 1-2*(34), 6–31.

Hermes, H., P. Lergetporer, F. Peter, and S. Wiederhold (2021). Behavioral Barriers and the Socioeconomic Gap in Child Care Enrollment. CESifo Working Paper Series 9282, Center for Economic Studies.

Hess, S. (2017). Randomization Inference with Stata: A Guide and Software. *Stata Journal 17*(3), 630–651.

Huber, M. and G. Mellace (2014). Testing Exclusion Restrictions and Additive Separability in Sample Selection Models. *Empirical Economics 47*(1), 75–92.

Huebener, M., A. Pape, and K. Spiess (2020). Parental Labour Supply Responses to the Abolition of Day Care Fees. *Journal of Economic Behavior and Organization 180*, 510–543.

Ilieva, B. and K. Wrohlich (2022). Gender Gaps in Employment, Working Hours and Wages in Germany: Trends and Developments Over the Last 35 Years. *CESifo Forum 23*(2), 17–19.

Jessen, J., S. Schmitz, and S. Waights (2020). Understanding Day Care Enrolment Gaps. *Journal of Public Economics 190*, 104252.

Kleven, H., C. Landais, J. Posch, A. Steinhauer, and J. Zweimüller (2019). Child Penalties Across Countries: Evidence and Explanations. *AEA Papers and Proceedings 109*, 122–26.

Kling, J. R., J. B. Liebman, and L. F. Katz (2007). Experimental Analysis of Neighborhood Effects. *Econometrica 75*(1), 83–119.

Ko, W. and R. A. Moffitt (2022). Take-up of social benefits. Working Paper 30148, National Bureau of Economic Research.

Kosse, F., T. Deckers, P. Pinger, H. Schildberg-Hoerisch, and A. Falk (2020). The Formation of Prosociality: Causal Evidence on the Role of Social Environment. *Journal of Political Economy 128*(2), 434–467.

Kuziemko, I., J. Pan, J. Shen, and E. Washington (2018). The Mommy Effect: Do Women Anticipate the Employment Effects of Motherhood? NBER Working Paper 24740, National Bureau of Economic Research.

Lee, D. S. (2009). Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects. *Review of Economic Studies 76*(3), 1071–1102.

List, J. A., A. M. Shaikh, and Y. Xu (2019). Multiple Hypothesis Testing in Experimental Economics. *Experimental Economics 22*(4), 773–793.

Lundberg, S. (2023). Gender Economics: Dead-Ends and New Opportunities. In *50th Celebratory Volume*, Volume 50 of *Research in Labor Economics*, pp. 151–189. Emerald Group Publishing Limited.

Müller, K. and K. Wrohlich (2020). Does Subsidized Care for Toddlers Increase Maternal Labor Supply? Evidence from a Large-Scale Expansion of Early Childcare. *Labour Economics 62*.

OECD (2011). *Doing Better for Families.* Paris: OECD.

OECD (2012). *Closing the Gender Gap: Act Now.* Paris: OECD.

OECD (2017a). *Dare to Share: Germany's Experience Promoting Equal Partnership in Families.* OECD, Paris.

OECD (2017b). *The Pursuit of Gender Equality: An Uphill Battle.* OECD, Paris.

OECD (2019). OECD Family Database: Enrolment in Childcare and Pre-School (PF3.2). `http://www.oecd.org/els/soc/pf3_2_enrolment_childcare_preschool.pdf`, OECD, Paris.

OECD (2020). OECD Family Database: Maternal Employment Rates. `https://www.oecd.org/els/family/lmf1_2_maternal_employment.pdf`, Paris.

OECD (2023). *Joining Forces for Gender Equality: What is Holding us Back?* OECD Publishing, Paris.

OECD (2023). OECD Family Database: Enrolment in Childcare and Pre-School (PF3.2). `http://www.oecd.org/els/soc/PF3_2_Enrolment_childcare_preschool.pdf`, OECD, Paris.

Oreopoulos, P., R. S. Brown, and A. M. Lavecchia (2017). Pathways to Education: An Integrated Approach to Helping At-Risk High School Students. *Journal of Political Economy 125*(4), 947–984.

Peichl, A., S. Sauer, and K. Wohlrabe (2022). Fachkräftemangel in Deutschland und Europa – Historie, Status quo und was getan werden muss. *ifo Schnelldienst 75*, 70–75.

Pora, P. and L. Wilner (2019). Child Penalties and Financial Incentives: Exploiting Variation along the Wage Distribution. Working Papers 2019-17, Center for Research in Economics and Statistics.

Ravazzini, L. (2018). Childcare and Maternal Part-Time Employment: A Natural Experiment Using Swiss Cantons. *Swiss Journal of Economics and Statistics 154*(15).

Resnjanskij, S., J. Ruhose, S. Wiederhold, L. Woessmann, and K. Wedel (2023). Can Mentoring Alleviate Family Disadvantage in Adolescence? A Field Experiment to Improve Labor-Market Prospects. *Journal of Political Economy Forthcoming*.

Romano, J. P. and M. Wolf (2005). Stepwise Multiple Testing as Formalized Data Snooping. *Econometrica 73*(4), 1237–1282.

Romano, J. P. and M. Wolf (2016). Efficient Computation of Adjusted P-Values for Resampling-Based Stepdown Multiple Testing. *Statistics & Probability Letters 113*, 38–40.

Schiman, C. (2022). Experimental evidence of the effect of head start on mothers' labor supply and human capital investments. *Review of Economics of the Household 20*(1), 199–241.

Scholz, A., K. Erhard, S. Hahn, and D. Harring (2018). Inequalities in Access to Early Childhood Education and Care in Germany. ICEC Working Paper Series, German Youth Institute (Deutsches Jugendinstitut ev.V.), Munich.

Silverman, B. (1986). *Density Estimation for Statistics and Data Analysis*. New York: Chapman & Hall.

SOEP (2019). Socio-Economic Panel (SOEP) 2019, Years 1984-2017, Version 34. `doi:10.5684/soep.v34`, DIW Berlin, Berlin.

Spiess, C. K. (2008). Early childhood education and care in Germany: The status quo and reform proposals. *Zeitschrift für Betriebswirtschaftslehre 2008 67*, 1–20.

Van Lancker, W. (2018). Reducing Inequality in Childcare Service Use across European Countries: What (if any) is the role of Social Spending? *Social Policy & Administration 52*(1), 271–292.

Van Lancker, W. and J. Ghysels (2016). Explaining patterns of inequality in childcare service use across 31 developed economies: A welfare state perspective. *International Journal of Comparative Sociology 57*(5), 310–337.

Walters, C. R. (2018). The Demand for Effective Charter Schools. *Journal of Political Economy 126*(6), 2179–2223.

Westfall, P. H. and S. S. Young (1993). *Resampling-Based Multiple Testing: Examples and Methods for P-Value Adjustment*. Wiley Series in Probability and Mathematical Statistics. New York: Wiley.

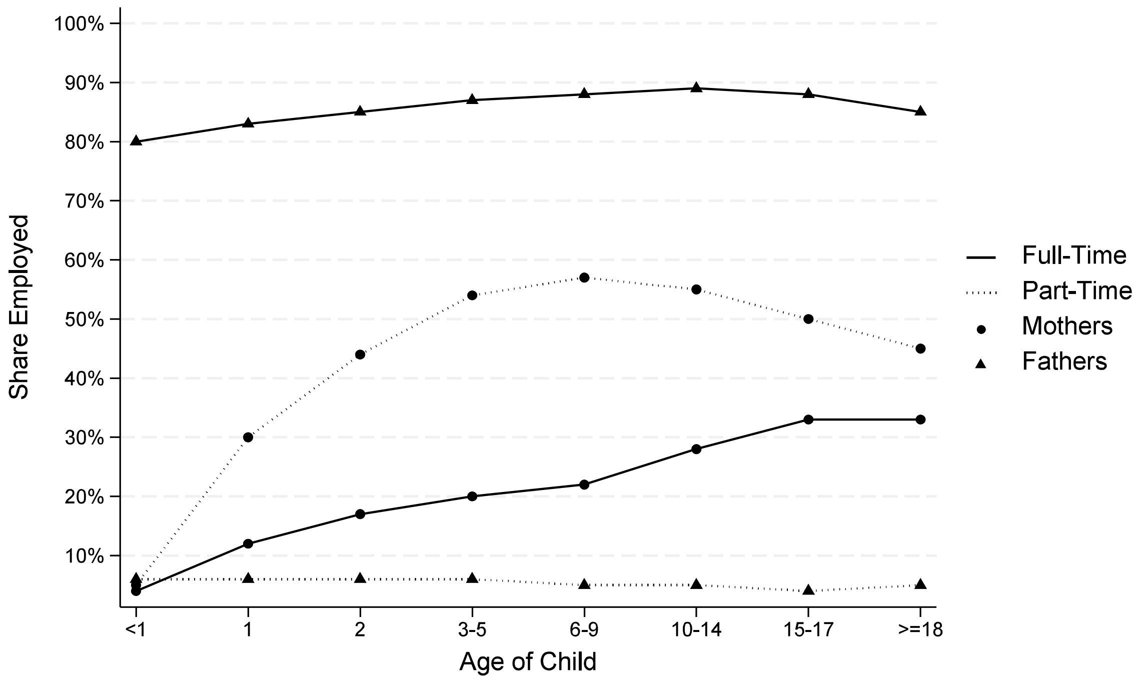
Wikle, J. and R. Wilson (2022). Access to Head Start and Maternal Labor Supply: Experimental and Quasi-Experimental Evidence. *Journal of Labor Economics forthcoming*.

Young, A. (2019). Channeling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results. *Quarterly Journal of Economics 134*(2), 557–598.

Online Appendix

Appendix A. Institutional Background

Figure A1: Full-Time and Part-Time Employment Rates of Mothers and Fathers



Notes: Figure is based on data from the German Microcensus from 2019 (BPB, 2021). Solid (dotted) lines depict the share of full-time (part-time) employed fathers (represented by triangles) and mothers (represented by dots). Reassuringly, the share of full-time-employed mothers with a child between 2 and 3 years is almost identical to the share in our sample (16.9% vs. 16.3%).

Appendix B. Further Details on Variables

Table B1: Variable Definitions

	Data Collection Date (1)	Definition (2)	Missings (3)	N (4)
Outcome variables				
Child care enrollment	9 months after treatment	Indicator of attending a child care center or having secured a slot in a child care center for the future, zero otherwise.	Missing if respondents did not participate in the first post-treatment survey ($n = 21$).	460
Full-time care (25h or more)	9 months after treatment	Indicator of the child spending 25 hours or more in a child care center in a typical week. If hours spent in child care are missing, we use values from the second post-treatment survey (only for those children enrolled in the first post-treatment survey).	Missing if respondents did not participate in the first post-treatment survey ($n = 21$).	460
Full-time care (30h or more)	9 months after treatment	Indicator of the child spending 30 hours or more in a child care center in a typical week. If hours spent in child care are missing, we use values from the second post-treatment survey (only for those children enrolled in the first post-treatment survey).	Missing if respondents did not participate in the first post-treatment survey ($n = 21$).	460
Hours in child care	9 months after treatment	Hours a child spent in a child care center in a typical week. If hours spent in child care in a typical week are missing, we use values from the second post-treatment survey (only for those children enrolled in the first post-treatment survey).	Missing if respondents did not participate in the first post-treatment survey ($n = 21$).	460
Hours with father	9 months after treatment	Care hours provided by the father alone in a typical week.	Missing if respondents did not participate in the first post-treatment survey ($n = 21$) or if respondents did not report weekly hours with the father ($n = 19$).	441
Hours with other caretakers	9 months after treatment	Care hours provided by other caretakers (i.e., grandparents, older siblings, other relatives, friends, nannies, or childminders) alone in a typical week.	Missing if respondents did not participate in the first post-treatment survey ($n = 21$) or if respondents did not report weekly hours with other caretakers ($n = 21$).	439
Full-time employment	18 months after treatment	Indicator of the mother working full-time. When eliciting information about maternal employment status, we set a threshold of 30 hours for full-time employment to rule out that answers were affected by different perceptions of the term "full-time employment". Self-employed mothers are defined as working full-time if they work 30 hours or more per week. In a few cases ($n = 19$), mothers classified themselves as working part-time, but reported 30 or more working hours per week. We classify these mothers as full-time employees. Results remain unchanged if we instead define these mothers as working part-time. In four cases, mothers indicated that they work full-time but reported less than 30 working hours per week. We also classify these mothers as full-time employees, but results do not change if we consider them as working part-time. In a few cases ($n = 35$) we elicit information about maternal full-time employment from the father. For these mothers we have no information about the exact number of working hours, but about the type of employment (full-time, part-time, not employed). Results are robust to excluding all interviews conducted with the father.	-	481

(continued on next page)

Table B1: Continued

	Data Collection Date (1)	Definition (2)	Missings (3)	N (4)
Log household income	18 months after treatment	Logarithmized monthly net household income.	Missing for $n = 22$ respondents, as $n = 9$ did not know their current household income and $n = 13$ did not report it.	459
Employment	18 months after treatment	Indicator of the mother working. We classify mothers as employed if they are in any type of paid employment (full-time, part-time or self-employed).	-	481
Working hours	18 months after treatment	Weekly working hours of working mothers according to contract. We recode the working hours variable to 30 hours per week in the four cases in which mothers indicated that they work full-time but reported less than 30 working hours per week. Results are robust if we use working hours as reported by the mothers.	Missing if mothers were not working ($n = 237$), if interviews were conducted with the father ($n = 11$), or if mothers did not indicate their working hours ($n = 10$).	223
Earnings	18 months after treatment	Monthly net earnings in the previous month. Earnings of marginally employed mothers (part-time jobs with earnings below 450 EUR) are calculated as monthly working hours times minimum wage in 2020 (i.e., 9.35 EUR).	Missing if mothers were not working ($n = 237$), if interviews were conducted with the father ($n = 11$), or if mothers did not indicate their earnings ($n = 10$).	223
Male breadwinner household	18 months after treatment	Households in which the father worked full-time and the mother worked at most part-time.	Missing for single mothers ($n = 40$).	441
Share maternal care hours	9 months after treatment	Calculated by dividing the care hours provided by the mother by the sum of both partners' care hours.	Missing if respondents did not participate in the first post-treatment survey ($n = 21$), if they did not answer the question about care hours ($n = 24$), or if they are single mothers ($n = 31$).	405
Share maternal earnings	18 months after treatment	Calculated by dividing maternal earnings by the sum of both partners' earnings. We have no direct information about fathers' earnings. Instead, we have information about maternal (net) earnings and (net) household income. We calculate fathers' earnings by subtracting maternal earnings from household income. With comparable information from the German Socio-Economic Panel (SOEP), we can benchmark the share of maternal earnings in parental earnings in our data with those from a nationwide representative sample. Reassuringly, our data are very much in line with the representative survey data. The share of maternal earnings in parental earnings in our total sample equals 0.33, being 0.28 for lower-SES mothers. In comparison, in the SOEP, the share of maternal earnings in parental earnings (for working mothers with children aged 2–3 years) amounts to 0.33 for the total sample and 0.25 for the lower-SES subsample.	Missing if mothers were not working ($n = 237$), if interviews were conducted with the father ($n = 11$), if mothers did not indicate their earnings or household income ($n = 12$), or if they are are single mothers ($n = 15$).	206
SES background				
Lower-SES	Baseline	As described in Section 3.2, we define mothers' SES background based on maternal education. Mothers with a college entrance qualification ("Abitur") are classified as higher-SES, whereas mothers without a college entrance qualification are classified as lower-SES. For three mother without education information, we use the partner's education level to assign SES background, based on the idea of assortative mating on education in Germany (Eika et al., 2019). Note that our results do not change if we exclude mothers with missing education information or classify all mothers with missing information as either lower-SES or higher-SES.	-	481

Table B1: Continued

	Data Collection Date (1)	Definition (2)	Missings (3)	N (4)
Control variables				
Pre-birth labor market outcomes				
Mother worked full-time (pre-birth)	Baseline	Indicator of full-time employment in the year before the child was born, indicated at baseline. Self-employed are classified as not working full-time (since their earnings are rather low on average and we did not elicit pre-birth working hours).	-	481
Mother worked (pre-birth)	Baseline	Indicator of employment (full-time or part-time) in the year before the child was born, indicated at baseline.	-	481
Mother's earnings (0 if not working, pre-birth)	Baseline	Mother's earnings in the year before the child was born, indicated at baseline. We assign zero earnings to mothers who did not work before the child was born.	-	481
Sociodemographic controls				
Mother is main caregiver	Baseline	Indicator of the mother being the main caregiver for the child.	-	481
Age of mother (in years)	Baseline	Age of mother in years.	-	481
Migration background	Baseline	Indicator of the the mother not being born in Germany.	-	481
Mother works or searches for a job	Baseline	Indicator of the mother having worked or having searched for a job at baseline.	-	481
Log household income	Baseline	Logarithmized monthly net household income.	-	481
Interview conducted with the father	Baseline	Indicator of the father having answered the baseline survey.	-	481
Controls mediation analysis				
Hours in child care	Baseline	Hours the child spent in a child care center in a typical week.	-	481
Hours with father	Baseline	Care hours provided by the father alone in a typical week.	-	481
Hours with other caretakers	Baseline	Care hours provided by other caretakers (i.e., grandparents, older siblings, other relatives, friends, nannies, or childminders) alone in a typical week.	-	481
Survey date fixed effects	Baseline	Fixed effects for the date of conducting the first post-treatment survey.	-	481
Controls gender equality				
Male breadwinner household (pre-birth)	Baseline	Households in which the father worked full-time and the mother worked at most part-time in the year before the child was born. As we have no information on fathers' labor supply before the child was born, we use fathers' labor supply at baseline as a proxy. The majority of fathers worked full-time at baseline (77%), which suggests that most fathers did not decrease their working hours substantially after childbirth. This is in line with evidence for Germany that mainly mothers take parental leave (Destatis, 2021).	-	481
Share maternal care hours	Baseline	Calculated by dividing the care hours provided by the mother by the sum of both partners' care hours.	-	481
Mother's earnings (0 if not working, pre-birth)	Baseline	See above. When estimating treatment effects on share maternal earnings, we use pre-birth earnings of the mother as control for the pre-treatment outcome because we do not have information on pre-birth earnings of the father.	-	481

Appendix C. Balancing and Attrition

Table C1: Baseline Sample: Descriptives and Balancing Tests

	All (1)	Control (2)	Treatment (3)	Δ(3)-(2) (4)	p-val for (4) (5)	p-val by SES (6)	Observations (7)
Pre-birth labor market outcomes							
Mother worked full-time	0.461	0.471	0.452	-0.020	0.635	0.929	588
Mother worked full-time, missing	0.031	0.031	0.031	-0.000	0.994	0.718	607
Mother worked	0.667	0.662	0.672	0.010	0.797	0.195	607
Mother's earnings (EUR)	1721	1724	1719	-5	0.963	0.231	386
Mother's earnings, missing	0.364	0.376	0.353	-0.023	0.554	0.217	607
Mother's earnings (0 if not working, in EUR)	1168	1156	1178	22	0.804	0.839	569
Mother's earnings (0 if not working), missing	0.063	0.070	0.056	-0.013	0.498	0.751	607
Sociodemographic characteristics							
Mother is main caregiver	0.967	0.976	0.959	-0.016	0.258	0.606	607
Age of mother (in years)	31.31	31.09	31.52	0.43	0.326	0.394	568
Migration background	0.390	0.389	0.391	0.002	0.953	0.831	568
Mother works or searches for a job	0.107	0.108	0.106	-0.002	0.944	0.320	607
Household income (EUR)	3050	2934	3156	222	0.123	0.082	576
Household income, missing	0.051	0.042	0.059	0.018	0.323	0.424	607
Interview conducted with the father	0.064	0.063	0.066	0.003	0.884	0.801	607
No school degree	0.054	0.052	0.056	0.004	0.829		607
Lower secondary degree	0.137	0.146	0.128	-0.018	0.516		607
Middle secondary degree	0.237	0.240	0.234	-0.006	0.862		607
College entrance degree	0.572	0.561	0.581	0.020	0.615		607

Notes: Table reports mean values of pre-birth labor market outcomes and sociodemographic characteristics in our baseline sample. All variables come from the baseline survey (i.e., before the treatment); labor market outcomes refer to one year before childbirth. Column (1) reports mean values for the full sample, Column (2) for the control group, and Column (3) for the treatment group. In Column (4), we show the difference between treatment and control groups, and Column (5) shows the corresponding p-value of a two-sided t-test testing the null hypothesis that values in Columns (2) and (3) are equal. In Column (6), we test whether there are treatment-control differences in the respective variable within SES subgroups. To conduct the test, we regress the variable on the treatment indicator, a higher-SES dummy, and their interaction. Column (6) reports the p-value of an F-test of joint significance of the coefficients on the treatment indicator and its interaction with the higher-SES dummy. *Mother worked full-time, pre-birth* is a dummy equal to one if the mother worked full-time before the child was born, zero otherwise. Information on maternal pre-birth work status was not reported in 19 cases (*Mother worked full-time, missing, pre-birth*). *Mother worked, pre-birth* is a dummy equal to one if the mother worked part-time or full-time pre-birth, zero otherwise. *Mother's earnings, pre-birth* denotes the monthly net earnings of the mother in EUR before the child was born. *Mother's earnings, missing, pre-birth* is a dummy equal to one if either the mother was not working pre-birth (183 cases) or the earnings information for the mother was not provided because the interview was conducted with the father (33 cases) or the mother did not answer the question (5 cases), zero otherwise. In *Mother's earnings (0 if not working), pre-birth* we assign zero earnings to mothers who did not work before the child was born; we use this variable when controlling for the pre-birth earnings of mothers. *Mother's earnings (0 if not working), missing, pre-birth* indicates cases in which maternal pre-birth earnings are missing because the father answered the baseline survey or because the mother did not answer the question. *Mother is main caregiver* is a dummy equal to one if the mother is the main caregiver of the child, zero otherwise. *Migration background* is a dummy equal to one if the mother was not born in Germany, zero otherwise. *Mother works or searches for a job* is a dummy equal to one if the mother worked at baseline (part-time or full-time) or searched for a job, zero otherwise. *Household income* is the monthly net household income in EUR. *No school degree*, *Lower secondary degree*, *Middle secondary degree* ("MSA"), and *College entrance qualification* ("Abitur") are all dummy variables indicating the parent's highest school degree.

Table C2: Check for Selective Attrition

	Participation Second Post-Treatment Survey			
	(1)	(2)	(3)	(4)
Treatment	-0.002	-0.010	0.006	-0.012
	(0.032)	(0.045)	(0.054)	(0.069)
Treatment × Mother Worked Full-Time, Pre-Birth		0.012		0.049
		(0.062)		(0.105)
Mother Worked Full-Time, Pre-Birth		0.012		0.078
		(0.046)		(0.077)
Treatment × Higher-SES			-0.015	-0.000
			(0.066)	(0.089)
Higher-SES			0.050	0.121*
			(0.050)	(0.063)
Treatment × Higher-SES × Mother Worked Full-Time, Pre-Birth				-0.049
				(0.129)
Higher-SES × Mother Worked Full-Time, Pre-Birth				-0.132
				(0.092)
Strata Controls	Yes	Yes	Yes	Yes
Sociodemographic Controls	Yes	Yes	Yes	Yes
N	607	607	607	607
Attrition F-test p-values				
Overall Sample	0.947	0.974	0.968	0.986
Lower-SES Sample			0.912	0.645
Higher-SES Sample			0.817	0.795

Notes: Table shows results from OLS models. Dependent variable is a dummy indicating participation in the second post-treatment survey. In Column (2) and (4), we impute missing values for *Mother worked full-time, pre-birth* with the median (zero) to not lose observations; results are qualitatively similar when we impute missings with one instead. All regressions include strata controls and baseline sociodemographic controls (see Section 4 for details). Imputation dummies for missing values in control variables are included. F-test p-values report p-values from joint significance tests of all treatment-related coefficients for the indicated sample. Robust standard errors in parentheses, * $p<.10$, ** $p<.05$, *** $p<.01$

Table C3: Treatment Effects on Full-Time Employment Using Inverse Probability Weighting

	Mother Works Full-Time		
	Unweighted	Baseline Sample Weights I	Baseline Sample Weights II
	(1)	(2)	(3)
Treatment	0.092**	0.091**	0.091**
	(0.043)	(0.042)	(0.041)
Treatment × Higher-SES	-0.097	-0.094	-0.094
	(0.064)	(0.062)	(0.062)
Higher-SES	0.110**	0.105**	0.106**
	(0.044)	(0.042)	(0.043)
Pre-Treatment Outcome	Yes	Yes	Yes
Strata Controls	Yes	Yes	Yes
Sociodemographic Controls	Yes	Yes	Yes
Treatment Effect	-0.005	-0.003	-0.004
Higher-SES	(0.047)	(0.045)	(0.046)
N	481	481	481

Notes: Table shows treatment effects using inverse probability weighting to account for attrition. The dependent variable is maternal full-time employment. Coefficients are weighted least squares estimates, with attrition weights being the inverse of the predicted probability of responding in the post-treatment survey. In Column (2), the probability of responding is derived from a probit model of the binary participation indicator as function of treatment assignment and pre-treatment outcome; in Column (3), the probability of responding is derived from a probit model of the binary participation indicator as function of treatment assignment, higher-SES indicator, their interaction, and pre-treatment outcome. All models include the pre-treatment outcome, strata controls, and baseline sociodemographic controls (see Section 4 for details). Imputation dummies for missing values in control variables are included. Robust standard errors in parentheses, * $p<.10$, ** $p<.05$, *** $p<.01$

Appendix D. Further Results

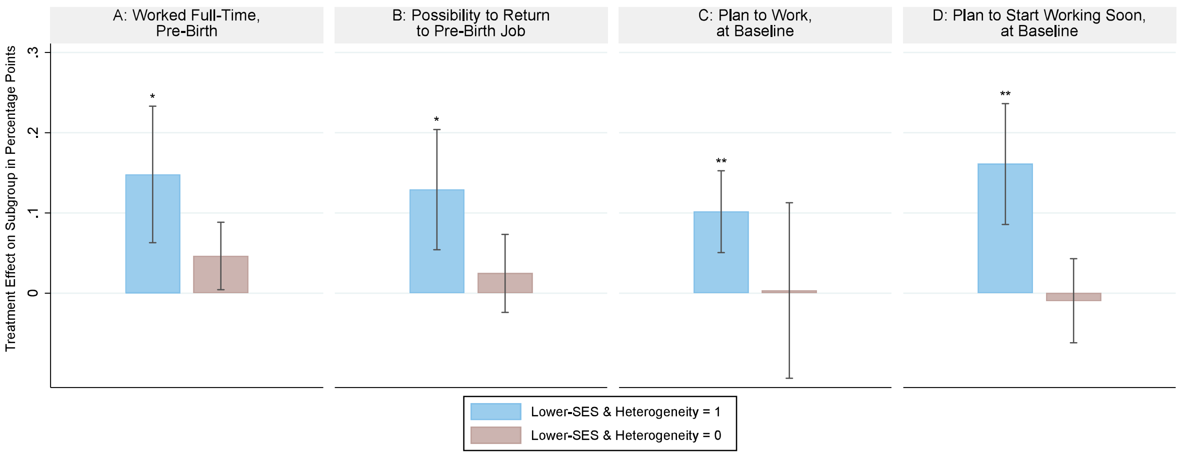
Heckman Selection Model. We restrict our sample to working mothers when estimating treatment effects on hours worked and earnings of mothers. Here, we use the Heckman selection model as a robustness check since it allows us to account for potential selection into the subsample of working mothers. In particular, the model first estimates the probability that a mother is employed (i.e., that she is observed in the sample of working mothers). The model then estimates the treatment effect on working hours and earnings while controlling for the fact that we did not observe all mothers in this sample. Results are reported in Table D1. As exclusion restrictions, we use either the existence of another child (born after baseline) (Columns (1) and (2)) or the total number of children in the household (Columns (3) and (4)) (Huber and Mellace, 2014). Results for both instruments are very similar.

Table D1: Treatment Effects on Weekly Working Hours and Monthly Earnings (Heckman Selection Model)

	Heckman Model			
	Working Hours	Log Earnings	Working Hours	Log Earnings
	(1)	(2)	(3)	(4)
Treatment	4.404**	0.221*	5.197**	0.232*
	(2.140)	(0.128)	(2.078)	(0.126)
Treatment × Higher-SES	-4.887*	-0.196	-6.550**	-0.226
	(2.699)	(0.172)	(2.608)	(0.166)
Higher-SES	4.059**	0.170	4.931**	0.197
	(2.062)	(0.132)	(2.002)	(0.136)
Pre-Treatment Outcome	Yes	Yes	Yes	Yes
Strata Controls	Yes	Yes	Yes	Yes
Sociodemographic Controls	Yes	Yes	Yes	Yes
Treatment Effect	-0.483	0.025	-1.353	0.007
Higher-SES	(1.659)	(0.106)	(1.584)	(0.103)
Control Mean Higher-SES	25.314	7.218	25.676	7.227
Control Mean Lower-SES	19.000	6.701	18.743	6.682
Control Mean SES Gap	6.314	0.517	6.934	0.545
Instrumental Variable	Younger Child	Younger Child	Number of Siblings	Number of Siblings
N	439	439	430	430

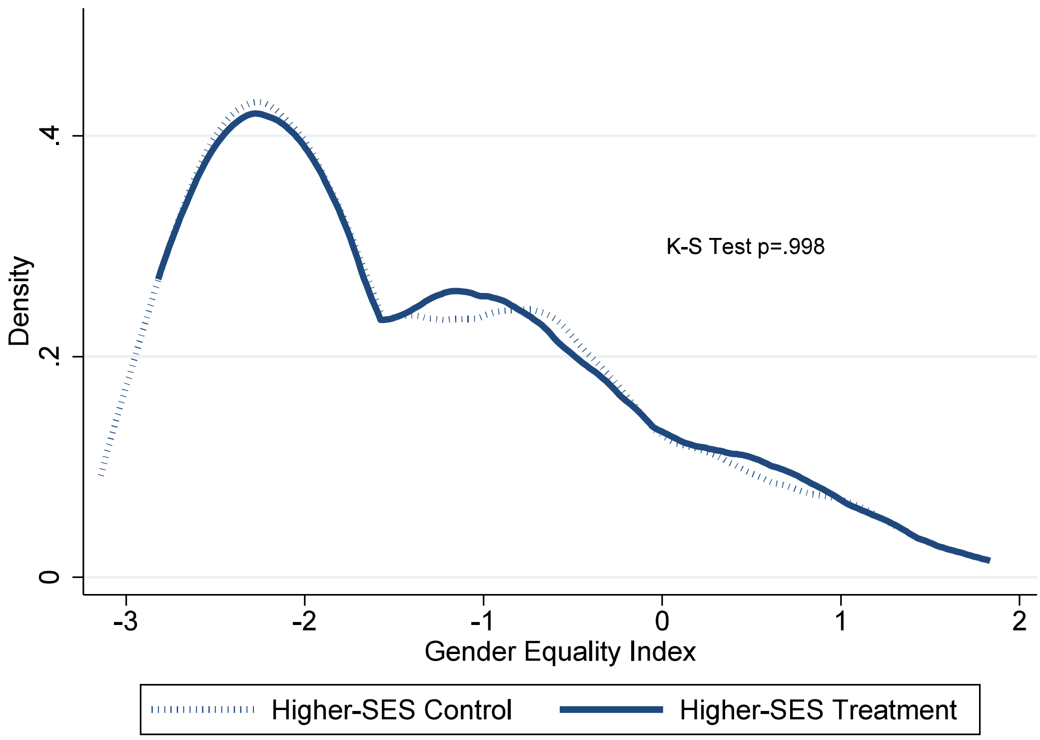
Notes: Table shows the second-stage results of a Heckman selection model for weekly working hours and log (net) earnings. Outcomes are measured when children are 2–3 years old (18 months after the treatment). The sample size deviates from our analytical sample, as we exclude families in which the father was the respondent (as we have no information about maternal working hours or wages in these cases). In Column (3) and (4), we additionally exclude 9 mothers because of missing information about the number of children in the household. In Column (1) and (2), we use the presence of a younger child as our targeted child as an instrument for selection into employment; in Column (3) and (4), the instrument is the number of siblings at baseline. *Higher-SES* equals one if the mother has a college entrance qualification, zero otherwise. All models include controls for pre-treatment outcomes (full-time employment of mother in the year before the child was born in Column (1) and (3), and earnings of mother in the year before the child was born in Column (2) and (4)). Additionally, we include controls for strata variables and baseline sociodemographic controls (see Section 4 for details). Imputation dummies for missing values in control variables are included. *Control Mean Higher-SES (Lower-SES)* is the mean of the respective outcome in the control group in the post-treatment survey for higher-SES (lower-SES) mothers; *Control Group SES Gap* reports the difference between control means of higher-SES and lower-SES mothers. Robust standard errors in parentheses, * $p<.10$, ** $p<.05$, *** $p<.01$.

Figure D1: Heterogeneous Treatment Effects on Maternal Full-Time Employment



Notes: Figure shows heterogeneous treatment effects (ITT) on maternal full-time employment for different subgroups within the lower-SES sample, based on OLS models using triple interactions. In Column (1), *Worked Full-Time, Pre-Birth* is equal to one if the mother worked full-time in the year before the child was born, zero otherwise ($n = 474$). *Worked Full-Time, Pre-Birth* is missing for $n = 7$ mothers because interviews were conducted with the father. Column (2) reports heterogeneity based on a dummy variable indicating whether the mother had the (perceived) possibility to return to her pre-birth job ($n = 461$), measured in the second post-treatment survey. *Possibility to Return to Pre-Birth Job* is missing for $n = 20$ mothers because interviews were conducted with the father. In Column (3), *Plan to Work, at Baseline* is equal to one if the mother indicates at baseline that she plans to work part-time or full-time, zero if she does not plan to take-up (paid) work ($n = 457$). *Plan to Work, at Baseline* is missing for $n = 24$ mothers, as $n = 19$ interviews were conducted with the father and $n = 5$ mothers did not answer the question. Column (4) reports heterogeneity based on whether the mother indicates at baseline that she plans to return to work in the next year (when the child is 1–2 years old) or earlier, zero if the mother plans to start working later or does not plan to take-up (paid) work ($n = 454$). *Plan to Start Working Soon, at Baseline* is missing for $n = 27$ mothers, as $n = 19$ interviews were conducted with the father and $n = 8$ mothers did not answer the question. Within each panel, the left-hand bar shows the estimated treatment effect for the subgroup of lower-SES mothers to which the respective heterogeneity applies (e.g., those mothers who worked full-time in the year before the child was born in Panel A); the right-hand bar shows the treatment effect for the remaining lower-SES mothers (e.g., those mothers who did not work full-time in the year before the child was born in Panel A). Full-time employment is measured when children are 2–3 years old (18 months after the treatment, see Section 4 for details). All models include the pre-treatment outcome, strata controls and baseline sociodemographic controls (see Section 4 for details). Imputation dummies for missing values in control variables are included. Error bars show robust standard errors. Significance levels: * $p< .10$, ** $p< .05$, *** $p< .01$.

Figure D2: Treatment Effect for Higher-SES Mothers on the Distribution of Gender Equality



Notes: Figure shows the unconditional treatment effect on the distribution of the gender equality index for higher-SES mothers. The dashed/dotted (solid) line shows the distribution for lower-SES mothers in the control (treatment) group. The probability density functions are computed with an Epanechnikov kernel with bandwidth h derived from the Silverman rule (Silverman, 1986) with $h = 0.9An^{-\frac{1}{5}}$, where n is the number of observations and $A = min(standard\ deviation, interquartile\ range/1.349)$. K-S Test: p-value for a Kolmogorov-Smirnov test of the equality of distributions of the gender equality index of treatment and control group.

Appendix E. Robustness Checks

Table E1: Treatment Effects on Full-Time Employment Using Full-Cohort Weights

	Mother Works Full-Time	
	Unweighted (1)	Full Cohort Weights (2)
Treatment	0.092**	0.090**
	(0.043)	(0.046)
Treatment × Higher-SES	-0.097	-0.112*
	(0.064)	(0.066)
Higher-SES	0.110**	0.114**
	(0.044)	(0.045)
Pre-Treatment Outcome	Yes	Yes
Strata Controls	Yes	Yes
Sociodemographic Controls	Yes	Yes
Treatment Effect	-0.005	-0.022
Higher-SES	(0.047)	(0.047)
Control Mean Higher-SES	0.234	0.228
Control Mean Lower-SES	0.056	0.061
Control Mean SES Gap	0.178	0.167
N	481	481

Notes: Table shows treatment effects using propensity score weights to account for selection into our study. Full-time employment is defined as working 30 hours or more per week and is measured when children are 2–3 years old (18 months after the treatment). The regression in Column (2) is re-weighted according to the full birth cohort in the two sample cities. Weights are derived from a probit model of a binary participation variable (indicating whether the family participated in the baseline survey) regressed on all variables available for the full cohort (age of parents and child, sex of child, child's migration background, child has German citizenship, child lives with both parents, zip-code and a proxy for first-time parent status and number of siblings). All models include the pre-treatment outcome, strata controls and baseline sociodemographic controls (see Section 4 for details). Imputation dummies for missing values in control variables are included. *Control Mean Higher-SES* (*Lower-SES*) is the mean of the outcome in the control group in the post-treatment survey for higher-SES (lower-SES) mothers; *Control Group SES Gap* reports the difference between control means of higher- and lower-SES mothers. In Columns (2) and (4), control means are re-weighted according to the full birth cohort. Robust standard errors in parentheses, * $p< .10$, ** $p< .05$, *** $p< .01$.

Table E2: Robustness Checks: Treatment Effect on Full-Time Employment of Mothers

	Mother Works Full-Time		
	Lockdown Indicator	Only Biological Mothers	Both Post-Treatment Surveys
	(1)	(2)	(3)
Treatment	0.089**	0.089*	0.077*
	(0.043)	(0.048)	(0.043)
Treatment × Higher-SES	-0.095	-0.108	-0.075
	(0.063)	(0.068)	(0.064)
Higher-SES	0.109**	0.108**	0.113**
	(0.044)	(0.047)	(0.046)
Pre-Treatment Outcome	Yes	Yes	Yes
Strata Controls	Yes	Yes	Yes
Mother characteristics	Yes	Yes	Yes
Treatment Effect	-0.006	-0.019	0.002
Higher-SES	(0.046)	(0.048)	(0.047)
Control Mean Higher-SES	0.234	0.244	0.230
Control Mean Lower-SES	0.056	0.061	0.049
Control Mean SES Gap	0.178	0.183	0.181
N	481	446	460

Notes: Table shows robustness checks for the intention-to-treat effect on full-time employment of mothers. In Column (1) we control for the interview being conducted during the first COVID-19 lockdown ($n = 481$). In Column (2) we estimate the treatment effect in the sample, in which the biological mother participated in the second post-treatment survey ($n = 446$) and in Column (3) we estimate the treatment effect in the subsample of families, who participated also in the first post-treatment survey ($n = 460$). All models are estimated by OLS. Full-time employment is defined as working 30 hours or more per week and is measured when children are 2–3 years old (18 months after the treatment, see Section 4 for details). *Higher-SES* equals one, if mother has a college entrance qualification. All models include the pre-treatment outcome, strata controls and baseline sociodemographic controls (see Section 4 for details). Imputation dummies for missing values in control variables are included. *Control Mean Higher-SES* (*Lower-SES*) is the mean of the outcome in the control group in the post-treatment survey for higher-SES (lower-SES) mothers; *Control Group SES Gap* reports the difference between control means of higher- and lower-SES mothers. Robust standard errors in parentheses, * $p< .10$, ** $p< .05$, *** $p< .01$.

Table E3: Randomization Inference and Corrections for Multiple Hypothesis Testing

	Coefficient (1)	Rand. Inference (2)	List-Shaikh-Xu (3)	Westfall-Young (4)	Romano-Wolf (5)
Panel A: Full-Time Employment (Table 3), Employment, and Working Hours (Table 4)					
Full-Time Employment Lower-SES	0.092**	**0.042**	**0.043**	**0.060**	**0.061**
Employment Lower-SES	−0.008	0.903	0.912	0.902	0.902
Working Hours Lower-SES	4.903**	**0.068**	**0.053**	**0.049**	**0.048**
Panel B: Household Income and Maternal Earnings (Table 4)					
Log Household Income Lower-SES	0.104*	**0.057**	0.119	0.133	0.132
Log Earnings Lower-SES	0.222*	**0.092**	**0.092**	0.133	0.132
Panel C: Gender Equality in the Household (Table 5)					
Male Breadwinner Household Lower-SES	−0.159**	**0.023**	**0.076**	**0.062**	**0.055**
Share Maternal Care Hours Lower-SES	−0.111**	**0.029**	**0.062**	**0.071**	**0.071**
Share Maternal Earnings Lower-SES	0.070*	**0.088**	0.111	**0.089**	0.266

Notes: Table shows p-values for our main results when using randomization inference and adjusting for multiple hypothesis testing. All p-values $<$ .10 are printed in **bold**. For comparison, Column (1) displays coefficients and significance stars representing p-values from robust standard errors (* $p<$ .10, ** $p<$ .05, *** $p<$ .01) as reported in the main tables. Randomization inference (RI) p-values in Column (2) are obtained from RI with 1,000 permutations, assigning the treatment status randomly within strata (using the Stata command 'ritest' by Hess, 2017). In Columns (3)–(5), we implement three different methods to correct for multiple hypothesis testing (controlling the family-wise error rates) using bootstrap resampling techniques. Column (3) uses the method by List et al. (2019), Column (4) the stepdown-approach by Westfall and Young (1993), and Column (5) the approach by Romano and Wolf (2005, 2016). The procedures by Westfall-Young (using the Stata command 'wyoung' by Julian Reif) and Romano-Wolf (using the Stata command 'rwolf' by Clarke et al. (2020)) account for the stratified randomization,i.e., bootstrap samples are selected within each stratum. In Panel A, we correct for the fact that we use three outcomes for labor market participation: full-time employment, employment, and working hours. In Panel B, we correct for the fact that we use two outcomes for earnings: household income and maternal earnings. In Panel C, we correct for the fact that we use three outcomes to measure intra-household gender equality: male breadwinner household, share maternal care hours, and share maternal earnings. Note that some corrected p-values are smaller than the original p-values because they are based on bootstrap methods. We do not report adjusted p-values for higher-SES mothers because of their insignificance in the main analysis. All control variables from the respective baseline specification are included.

Table E4: Alternative SES Definitions

	Mother Works Full-Time			
	SES-1 (main)	SES-2	SES-3	SES-4
Treatment	0.092**	0.089*	0.070*	0.073*
	(0.043)	(0.051)	(0.037)	(0.037)
Higher-SES-1 (mother has a college entrance qualification)	0.110**			
	(0.044)			
Treatment × Higher-SES-1 (main)	-0.097			
	(0.064)			
Higher-SES-2 (at least one parent in HH has college entrance qualification)		0.056		
		(0.044)		
Treatment × Higher-SES-2		-0.080		
		(0.066)		
Higher-SES-3 (mother has college entrance qualification + HH not poor)			0.114**	
			(0.051)	
Treatment × Higher-SES-3			-0.077	
			(0.066)	
Higher-SES-4 (mother has college entr. qual. + no single mother + HH not poor)				0.112**
				(0.051)
Treatment × Higher-SES-4				-0.082
				(0.066)
Pre-Treatment Outcome	Yes	Yes	Yes	Yes
Strata Controls	Yes	Yes	Yes	Yes
Mother characteristics	Yes	Yes	Yes	Yes
Control Mean Higher-SES	0.234	0.209	0.275	0.275
Control Mean Lower-SES	0.056	0.058	0.059	0.059
Control Mean SES Gap	0.178	0.151	0.216	0.216
N	481	481	481	481

Notes: Table shows our main result from Table 3 for different definitions of lower-SES mothers (see Section 6.3 for details). Column (1) reports results for our main definition of lower-SES mothers used throughout the paper (mothers not having a college entrance qualification, SES-1). In Column (2), we define lower-SES as *neither* parent having a college entrance qualification (SES-2, 30% of the full sample). Column (3) extends the definition for SES-1 by adding families with an equivalent household income below poverty line to the lower-SES sample (SES-3, 51% of the full sample, following Falk et al. (2021)). Column (4) extends the definition for SES-3 by adding single-mothers to the lower-SES sample (SES-4, 52% of the full sample, following Kosse et al. (2020)). All models include the pre-treatment outcome, strata controls, and baseline sociodemographic controls (see Section 4 for details). Imputation dummies for missing values in control variables are included. *Control Mean Higher-SES* (*Lower-SES*) is the mean of the outcome in the control group in the post-treatment survey for higher-SES (lower-SES) mothers; *Control Group SES Gap* reports the difference between control means of higher- and lower-SES mothers. Robust standard errors in parentheses, * $p<.10$, ** $p<.05$, *** $p<.01$

Appendix F. Details on the Mediation Analysis

In our mediation analysis, we follow the approach by Heckman et al. (2013) and Heckman and Pinto (2015). Thus, we add the mediators to Equation (1), yielding:

$$\begin{aligned} Y_i = \alpha + \beta_1^{residual} Treatment_i + \beta_2^{residual} Treatment_i \times HigherSES_i \\ + \beta_3 HigherSES_i + \sum_k \theta^k M_i^k + \sum_k \lambda^k M_i^k \times HigherSES_i + \mathbf{X_i'}\delta + \mu_i \quad \text{(F.1)} \end{aligned}$$

We assume that the outcome is a linear function of our observed k mediators (M_i^k) and a vector of sociodemographic controls (X_i'). As our treatment effect varies by SES, we also allow effects of mediators to differ by SES (θ^k for lower-SES mothers and $\theta^k+\lambda^k$ for higher-SES mothers). Underlying this mediation analysis are three assumptions. First, as implied by Equation (F.1) we assume that the θ^ks do not differ between treatment and control groups. More precisely, this means that our included mediators (i.e., hours in child care, hours with the father, and hours with other caretakers) affect maternal full-time employment similarly in both treatment and control groups, conditional on our other mediators and control variables. Second, the influence of the control variables included in $\mathbf{X_i'}$ on full-time employment has to be similar in treatment and control groups. We test both assumptions by interacting our mediators (first assumption) and control variables (second assumption) with the treatment dummy. Reassuringly, no interaction term is statistically significant at the 5% level, while only hours in child care is significant at the 10% level.[45]

Finally, we assume that any unobserved mediator (included in the error term μ_i) that is affected by the treatment is orthogonal to the observed mediators. Otherwise, this would bias the estimated share of the treatment effect explained by our mediators in Figure 3. Given that these assumptions hold, $\beta_1^{residual}$ indicates the treatment effect on lower-SES mothers net of the observed mediators, and the share of the treatment effect explained by all observed mediators is $1 - \beta_1^{residual}/\beta_1$ (with β_1 from Equation (1)).

In addition, to assess the relative contribution of each mediator, we use the estimate of the treatment effect on the respective mediator (see Appendix Table F1):

$$M_i^k = \gamma_0^k + \gamma_1^k Treatment_i + \gamma_2^k Treatment_i \times HigherSES_i + \gamma_3^k HigherSES_i + \mathbf{X_i'}\boldsymbol{\gamma_4^k} + \eta_i^k \quad \text{(F.2)}$$

[45] We do not control for pre-treatment outcomes in the mediation analysis to ensure that the underlying assumptions are fulfilled.

The share of the treatment effect for lower-SES mothers attributed to mediator M^k can be expressed as $m_k = \theta^k \gamma_1^k / \beta_1$, with θ^k estimated from Equation (F.1), γ_1^k from Equation (F.2), and β_1 from Equation (1). Results are shown in Figure 3 and Appendix Table F2.

Table F1: Treatment Effects on Potential Mediators (Child Age 1–2Y)

	Non-Maternal Care Hours		
	Hours in Child Care	Hours with Father	Hours with Other Caretakers
	(1)	(2)	(3)
Treatment	4.009**	2.554*	-2.320*
	(2.027)	(1.424)	(1.407)
Treatment × Higher-SES	-5.460**	-1.965	1.435
	(2.694)	(1.880)	(1.871)
Higher-SES	6.767***	0.201	-1.483
	(2.074)	(1.139)	(1.442)
Pre-Treatment Outcome	Yes	Yes	Yes
Strata Controls	Yes	Yes	Yes
Sociodemographic Controls	Yes	Yes	Yes
Treatment Effect	-1.452	0.589	-0.886
Higher-SES	(1.857)	(1.190)	(1.203)
Control Mean Higher-SES	13.674	4.706	4.688
Control Mean Lower-SES	5.244	3.500	5.487
Control Mean SES Gap	8.430	1.206	-0.800
N	460	441	439

Notes: Table shows intention-to-treat effects on various types of non-maternal care hours, all models are estimated by OLS. Outcomes are measured when children are 1–2 years old (nine months after the treatment). In Column (1), the outcome is weekly hours in center-based child care. In Column (2), the outcome is weekly care hours provided by the father alone. In Column (3), the outcome is weekly hours in other (non-maternal) care arrangements (i.e., care hours provided by grandparents, older siblings, other relatives, friends, nannies, or childminders alone). Of the total sample of $n = 481$ mothers, $n = 21$ did not participate in the first post-treatment survey in which the outcomes were elicited, and for a few mothers we have no information on the respective outcome variable ($n = 19$ in Column (2), $n = 21$ in Column (3); for details, see Appendix B). All models include pre-treatment hours in child care, with the father and with other caretakers. In addition, they include strata controls, baseline sociodemographic controls, and survey date fixed effects (see Section 4 for details). Imputation dummies for missing values in control variables are included. *Control Mean Higher-SES* (*Lower-SES*) is the mean of the outcome in the control group in the post-treatment survey for higher-SES (lower-SES) mothers; *Control Group SES Gap* reports the difference between control means of higher- and lower-SES mothers. Robust standard errors in parentheses, * $p< .10$, ** $p< .05$, *** $p< .01$.

Table F2: Treatment Effects on Maternal Full-Time Employment with Mediators

	Main	Channel 1: Hours in Child Care	Channel 2: Hours with Father	Channel 3: Hours with Other Caretakers	All Channels
	(1)	(2)	(3)	(4)	(5)
Treatment	0.091*	0.066	0.078	0.083	0.041
	(0.054)	(0.051)	(0.053)	(0.055)	(0.049)
Treatment × Higher-SES	-0.104	-0.066	-0.092	-0.097	-0.040
	(0.074)	(0.070)	(0.073)	(0.075)	(0.068)
Higher-SES	0.137***	0.064	0.141**	0.106*	0.032
	(0.052)	(0.053)	(0.056)	(0.056)	(0.058)
Hours in Child Care		0.006*			0.006**
		(0.003)			(0.003)
Hours with Father			0.006		0.007*
			(0.004)		(0.004)
Hours with Other Caretakers				-0.003	-0.003
				(0.003)	(0.003)
Control Mean Higher-SES					0.232
Control Mean Lower-SES					0.050
Control Mean SES Gap					0.182
N	436	436	436	436	436

Notes: Table shows the models estimated from Equation (F.1), which are the basis for calculating the share of the treatment effect explained by specific mediators in our mediation analysis (see Section 6.4). We consider the following mediators: *Hours in Child Care* refer to weekly hours in center-based child care; *Hours with Father* refer to hours in which the father takes care of the child alone; *Hours with Other Caretakers* refer to hours in which other caretakers (i.e., grandparents, older siblings, other relatives, friends, nannies, or childminders) take care of the child alone. Column (1) replicates the main treatment effect on maternal full-time employment from Column (4) of Table 3, using the sample for which we have information on all mediators (n = 436). In Columns (2)–(4), we include each mediator separately (corresponding to the first three bars in Figure 3). In Column (5), we include all three mediators jointly (corresponding to the last bar of Figure 3). In all specifications, we also include interaction terms of the mediator variable with the higher-SES dummy, allowing the mediators to differently affect full-time employment for lower- and higher-SES mothers. All models include strata controls, baseline sociodemographic controls, and survey date fixed effects (see Section 4 for details). Imputation dummies for missing values in control variables are included. *Control Mean Higher-SES* (*Lower-SES*) is the mean of the outcome in the control group in the post-treatment survey for higher-SES (lower-SES) mothers; *Control Group SES Gap* reports the difference between control means of higher- and lower-SES mothers. Robust standard errors in parentheses, * $p<.10$, ** $p<.05$, *** $p<.01$.

Appendix G. Details on the Back-of-the-Envelope Calculation

We conduct a back-of-the-envelope calculation that compares the treatment-induced additional tax revenue (resulting from the additional income of mothers) with the costs for the provision of the additional hours in child care. We do not take the costs of the actual treatment into account for this calculation, as this cost is negligible (video and personal assistance provided by an RA of on average 2 hours per family).

Tax Revenue. As described in Section 6.2 and Table 4, the net household income of treated lower-SES mothers increases by 10.4%, which corresponds to 300 EUR (the average household income for lower-SES families in the control group in the second post-treatment survey is about 2900 EUR). Conservatively assuming a very low marginal income tax rate of 24% (the second lowest marginal tax rate in Germany), the additional tax revenue amounts to 95 EUR per month.

Child Care Costs. As shown in Section 5 and Table 2, hours in child care increase by four hours per week (or about 16 hours per month) due to our treatment. The provision of a full-time slot in child care costs about 850-1000 EUR per month (see, e.g., Felfe et al., 2015; Destatis, 2021), the price per hour amounts to about 6 EUR.[46] Therefore, the provision of 16 additional hours per month in child care would cost less than 100 EUR per child.

This simple calculation suggests that the treatment-induced increase in tax revenue due to the higher household income of lower-SES mothers would be sufficient to fully fund the necessary increase in child care provision.

[46] A full-time slot in child care is covering 35 hours per week, and with 900 EUR per month, we arrive at an hourly cost of about 6 EUR. Note that the number of hours for the cost calculation differs from the hours that parents actually *use* (see Section 5) because in the cost estimation we have to rely on the number of hours the child care center has to (potentially) provide for each child.

Appendix H. Experimental Material

In this section, we provide additional details on the material used for recruitment and panel maintenance, as well as details about the treatment video.

Appendix H.1. Material used to Communicate with Parents

Figure H1: Invitation Letter, Examples of Materials for Parents, and Study Logo

Notes: Figure presents material used in the study, including letters sent to parents and various postcards used for panel maintenance (letters are blurred for data protection reasons).

Appendix H.2. Treatment Video (Slides and Text)

Figure H2: Slide 1 of the Information Video Shown to the Treatment Group

Audio Slide 1: Who should take care of my child and where? Almost all parents with young children in Germany face these questions.

Figure H3: Slide 2 of the Information Video Shown to the Treatment Group

Audio Slide 2: This short video summarizes the most important information about child care for you. There are many different ways to care for your child. You can choose! Basically, you can look after your child yourself at home, your child could attend a child care center, or, for example, a nanny could take care of your child. The decision is entirely up to you.

Figure H4: Slide 3 of the Information Video Shown to the Treatment Group

Audio Slide 3: Access to a child care slot.

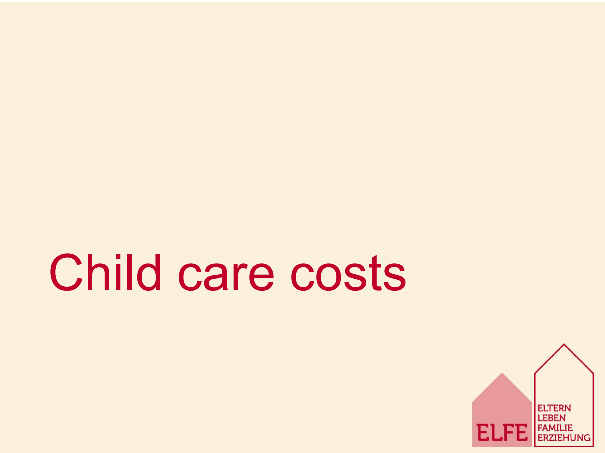
Figure H5: Slide 4 of the Information Video Shown to the Treatment Group

Audio Slide 4: Many parents think that their child cannot attend child care because there are no slots available or it is too expensive. But is this really the case? In Germany, all parents have a legal entitlement to a child care slot for their children from their first birthday onward. This applies without exceptions. And the legal entitlement also applies regardless of whether parents work or not.

Figure H6: Slide 5 of the Information Video Shown to the Treatment Group

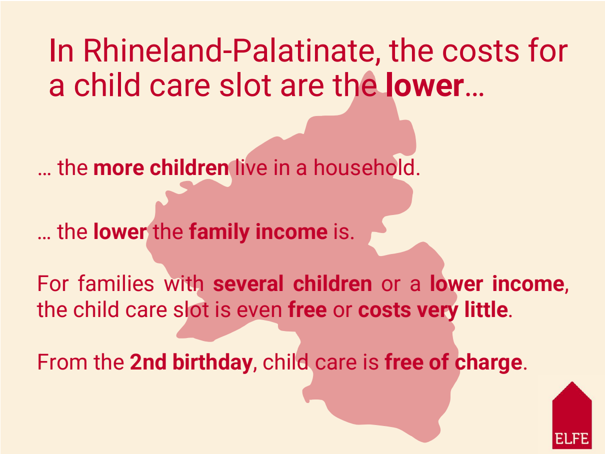
Audio Slide 5: You can choose the child care slot yourself. If you cannot find a slot in a child care center of your choice, the following applies: your city must offer you an alternative slot in child care, for example, at another child care center.

Figure H7: Slide 6 of the Information Video Shown to the Treatment Group



Audio Slide 6: Child care costs: Many parents think that a child care slot is very expensive. But how much does a child care slot really cost?

Figure H8: Slide 7 of the Information Video Shown to the Treatment Group



Audio Slide 7: In Rhineland-Palatinate, the costs for a child care slot are the lower the more children live in a household and the lower the family income is. For families with several children or a lower household income, child care is often even free or costs very little. From a child's second birthday, child care is even free of charge.

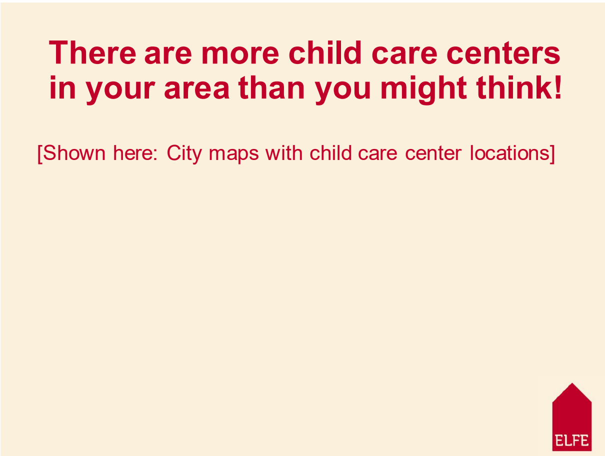
Figure H9: Slide 8 of the Information Video Shown to the Treatment Group

Audio Slide 8: Application for a child care slot.

Figure H10: Slide 9 of the Information Video Shown to the Treatment Group

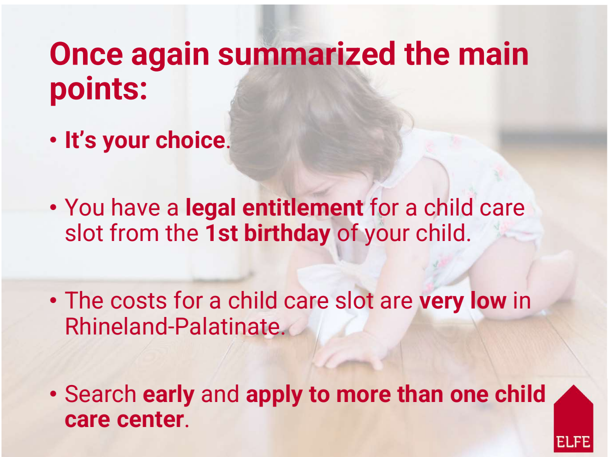
Audio Slide 9: If you decide to apply for child care, finding a slot is really not rocket science. The application procedures for a child care slot may differ from city to city. However, the following always applies: apply early. This increases your chances of finding a slot. In many child care centers, there is an application deadline. But even if there is no application deadline, the earlier you apply, the better. In any case, apply to more than one child care center! This will increase your chances of getting a slot at the child care center of your choice.

Figure H11: Slide 10 of the Information Video Shown to the Treatment Group



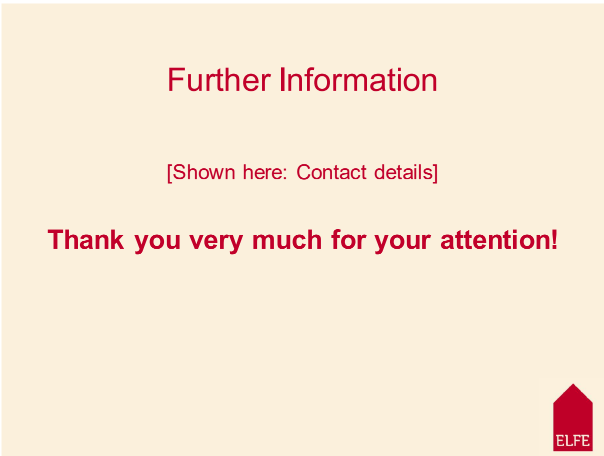
Audio Slide 10: Here you can see that there are more child care centers in your area than you might think!

Figure H12: Slide 11 of the Information Video Shown to the Treatment Group



Audio Slide 11: Once again, we summarize the main points. It's your choice whether you want to care for your child yourself at home or whether you want to enroll your child in child care. You have a legal entitlement to child care from the day your child is one year old. The costs for child care are very low in Rhineland-Palatinate, and, from the day your child is two years old, child care is even free of charge. If you would like to enroll your child in child care, search early and apply to more than one child care center. If you have any questions or need support regarding child care, please contact our staff. We will gladly help you!

Figure H13: Slide 12 of the Information Video Shown to the Treatment Group



Audio Slide 12: Further information can also be found at [webpage]. Thank you very much for your attention! We wish you and your family all the best and thank you for participating in the ELFE study.