The Interaction Effects of Subsidized Childcare and Parental Leave Entitlements on Labor Market Outcomes: Evidence from Germany

Hans Schwarz

Job Market Paper

This Version: April 24, 2022

Click here for the latest version

Abstract

I analyze the interaction effects of subsidized childcare and long job-protected parental leave on the duration of maternal leave and subsequent labor market outcomes. My identification strategy exploits the staggered roll-out of a federal expansion in the number of childcare centers for children ages 0–3 in Germany. Using the KiBS household survey and a generalized difference-in-differences approach, I find that an additional daycare center in a locality reduces the duration of maternal leave by more than 13 days (2.5 percent), which indicates the two family-friendly policies are substitutes. I then exploit the roll-out to examine heterogeneous treatment effects of taking a short maternal leave rather than a long leave on subsequent labor market outcomes. Using the Marginal Treatment Effects framework, I show that the take-up of a short maternal leave instead of a long leave has positive causal effects on maternal labor force participation and earnings after the child turns three for all mothers. Nonetheless, I find that the relative returns of a short leave are particularly high for the mothers that are the least likely to take a short leave. This implies that promoting a faster return to work of hard-to-reach mothers can have substantial effects on their subsequent labor market outcomes.

JEL Classification: J13, J16, J18, J21, J24

Keywords: Family-friendly policies, childcare, parental leave, labor market outcomes

*I am deeply grateful to my advisors John Kennan, Jeff Smith, and Chris Taber for their guidance and advice. I also want to especially thank Sandra Spirovsk a and Elise Marifian, as well as Md Moshi Ul Alam, Nicolas Badaracco, Jason Fletcher, Chao Fu, Renata Gain eddenova, Amrita Kulka, Lois Miller, Corina Mommaerts, Joseph Mullins, Minseon Park, Arpita Patnaik, Elan Segarra, Jim Walker, Matt Wiswall, and Andrey Zubanov, as well as seminar participants at the University of Wisconsin-Madison for their helpful comments. Finally, I am also indebted to Susanne Kuger and Anton Jeffrey of the German Youth Institute (DJI) for sharing restricted household location data from the German Child Care Study (KiBS) to use for the purposes of this paper. All errors are my own.

†Department of Economics, University of Wisconsin-Madison. Email: schwarz4@wisc.edu
1 Introduction

The onset of parenthood is a critical period for families in multiple ways. However, the effects of parenthood are unevenly distributed within the household. Recent research has highlighted that mothers carry a significantly higher share of the labor market penalties associated with parenthood than fathers (Angelov et al., 2016; Bertrand et al., 2010; Cortes and Pan, 2020; Kleven et al., 2019a,b).

In recent decades, governments around the world have become increasingly invested in the design and implementation of family-friendly policies that can boost declining fertility rates and narrow the gendered labor market gaps associated with childbirth. In most countries, policy makers have focused on expanding the availability of subsidized childcare and providing paid or unpaid job-protected leave to parents as key tools to promote work-life balance.\(^1\)

Many papers have previously studied the effects of expanding subsidized childcare or modifying different components of the parental leave system in isolation. The literature has found overall positive but small effects of subsidized childcare on maternal labor market outcomes (Olivetti and Petrongolo, 2017). In contrast, an often-cited consensus about the effects of maternity leave on future labor market outcomes is that short leave entitlements can have positive effects on subsequent outcomes, but that long leave entitlements appear to have negative effects (Olivetti and Petrongolo, 2017; Rossin-Slater, 2017).\(^2\)

In practice, however, the provision of free or subsidized childcare and the entitlement to paid or unpaid job-protected leave have similar policy goals. The two policies are most likely interdependent. For instance, the expected return of taking a short leave instead of a long leave after the birth of a child might crucially depend on how uncertain the parent is about securing a spot in subsidized childcare. A long leave might provide a valuable source of insurance against this risk, which might be substantial if the seats in subsidized childcare are rationed.

Importantly, the presence of close policy substitutes or complements can affect the cost-benefit assessment of a given policy (Kline and Walters, 2016; Johnson and Jackson, 2019; Bailey et al., 2020). Although this has been hypothesized in the literature for the case of parental leave entitlements and subsidized childcare, there is still very little empirical evidence about how these two policies jointly affect household decisions.\(^3\) The estimation of the interaction effects of these two family-friendly policies on subsequent labor market outcomes is an important first step towards a comprehensive assessment of the relative advantages and disadvantages of the two policies.

\(^{1}\)The U.S. is an outlier among the set of developed countries. Until very recently, the discussion about the role of government intervention in the promotion of family-friendly policies in the U.S. was centered almost exclusively on in-work benefits, which includes the Earned Income Tax Credit (EITC).

\(^{2}\)However, an opposite result has been found in the U.S. Bailey et al. (2019) shows that the introduction of six weeks of partially paid leave in California had negative long-run effects in the employment and wages of first-time mothers.

\(^{3}\)An exception is Danzer et al. (forthcoming), which analyzes the heterogeneous effects of extending the duration of paid leave in Austria on children’s outcomes, depending on whether the county had a daycare center for children ages 0 to 3 or not at the time of the leave extension. In a related note, Lalive et al. (2014) analyze the interaction between the job-protected leave and parental allowance components of the parental leave system in Austria. The authors conclude that the two are complements. Increasing only one component but not the other one does not substantially affect the return-to-work behavior of mothers. Finally, Wang (2019) analyzes the optimal design of parental leave and subsidized childcare in order to boost fertility rates and maternal labor market outcomes in Germany using a dynamic discrete choice model. Wang (2019) finds that offering higher childcare subsidies is a cost-effective way of increasing welfare.
Furthermore, the underlying causal mechanisms behind the empirical consensus that long leave entitlements are associated with worse labor market outcomes for parents are not completely understood. In particular, it is still an open question in the literature the extent to which this empirical result is driven by selection, where parents with different observed and unobserved characteristics self-select into different leave durations, or human capital depreciation.

The aim of this paper is two-fold. First, I analyze whether parental leave decisions are affected by the local availability of subsidized childcare. I study this interaction in the German context, where each parent is entitled to take up to three years of job-protected leave. In order to identify the average treatment effects of modifying the local availability of subsidized childcare on parental leave decisions, I exploit plausibly exogenous variation in the local availability of childcare stemming from the staggered implementation of an ambitious federal expansion in the number of childcare facilities and seats for children in the age group 0 - 3. Nationwide, the proportion of children ages 0 - 3 that attended childcare passed from 15.5% in 2007 to 32.3% in 2014 (Bauernschuster et al., 2016; Felfe and Lalive, 2018a; Müller and Wrohlich, 2020).

I use data from the German Child Care Study (KiBS), which is a big, representative sample of households with small children in Germany, for the analysis. A crucial advantage of the KiBS dataset over other datasets is that it contains retrospective information about parental leave durations.

Using a generalized difference-in-differences approach similar to the one first used in Berlinski and Galiani (2007), I estimate that one additional childcare center per 1,000 children ages 0–3 in a locality reduces the duration of maternal leave by slightly more than 13 days on average, which is equivalent to a 2.5% reduction in the mean duration of maternal leave in the sample. Using the same generalized difference-in-differences approach with an alternative discretized outcome that classifies leaves into “short leaves” and “long leaves” based on the 14-month window in which parents are eligible to receive any federal parental allowance, I find that one additional childcare center per 1,000 children ages 0–3 in the locality reduces the probability that a mother takes a long leave by 1.6 percentage points on average. Both results are robust to different specification and validation checks.

Another important finding of the paper is that paternal leave decisions do not seem to react to changes in the availability of subsidized childcare in the sub-sample of intact households. This result is not driven by lack of statistical power: I can rule out that an additional childcare center in the locality (per 1,000 children ages 0–3) increases the duration of paternal leave by more than 3 days at the 5 percent significance level. A similar result for a different set of employment outcomes is found in Norway (Andresen and Havnes, 2019). More generally, this empirical pattern is in line with papers that have concluded that the labor supply of married women is more elastic to exogenous shocks in the environment than the labor supply of married men (e.g., Blundell et al. (2016)).

In the second part of the paper, I proceed with the analysis of the heterogeneous treatment effects of maternal leave decisions on labor market outcomes of mothers after the child turns three years old. I accomplish this by using the Marginal Treatment Effects (MTE) framework developed

The MTE framework allows the presence of unrestricted empirical patterns of selection into treatment as a function of observed and unobserved returns to treatment. For example, the MTE framework allows for Roy-type positive (negative) selection into treatment, where the individuals that are the most likely to select into treatment are the ones who have on average the highest (lowest) returns from treatment. Nonlinear relationships are also admissible (Brinch et al., 2017; Carneiro et al., 2011; Cornelissen et al., 2018). In contrast, other structural approaches require more assumptions about the allocation of childcare seats that might be limiting the set of admissable relationships between returns and selection into treatment (Haan and Wrohlich, 2011; Geyer et al., 2015; Bick, 2016; Wang, 2019).

In my main MTE specification, I define the relevant treatment to be a binary variable on whether the mother takes a “short leave” rather than a “long leave” after the birth of the child. The nonparametric identification of the MTE is guaranteed with a continuous instrument that is independent of unobserved household characteristics that influence selection into treatment and outcomes in the treated or untreated state, after controlling for observed characteristics (Heckman and Vytlacil, 2005). I use the local availability of childcare, measured as the number of childcare centers that admit children in the age group 0–3, as my preferred instrument. Following the generalized difference-in-differences strategy, locality fixed effects and cohort fixed effects are also included in the vector of observed characteristics, similar to the strategy of Cornelissen et al. (2018).

I analyze the heterogenous treatment effects of maternal leave decisions on four different labor market outcomes: an indicator on whether the mother is working, the number of hours worked, the monthly maternal income, and an indicator variable on whether the household is receiving transfer payments. These outcomes are measured after the child turns three years old, when the child becomes eligible to attend kindergarten. To my knowledge, this is the first paper that estimates the causal effects of different leave durations on subsequent labor market outcomes using MTE.

I find that short maternal leaves relative to long maternal leaves have positive causal effects on subsequent labor market outcomes for all mothers. For example, I estimate that the average treatment effect (ATE) of a short leave relative to a long leave on hours worked after the child turns three is equal to 14.69 additional hours of work (s.e. 6.18) for mothers that gave birth between August of 2007 and July of 2014. The MTE is also positive throughout the common support of the unobserved resistance to treatment. The estimates of the MTE of short leaves on the probability of work and monthly maternal income are also positive throughout albeit noisier. In particular, the estimates of the ATE on the work indicator variable and on monthly income of the mother are 0.21 (s.e. 0.19) and €779.54 (s.e. €502.80), respectively.

My results also provide evidence that the effects are highly heterogeneous in terms of observed and unobserved characteristics. In particular, I find that the relationship between the unobserved resistance to treatment and the returns to treatment based on unobservable characteristics has a statistically significant u-shape form for all outcomes except transfer payment receipt. This means that the mothers that are either the most likely or the least likely to select into short leaves in the
sample have the highest unobserved returns to treatment on subsequent labor market outcomes. In the presence of heterogeneous treatment effects with a binary instrument, Two-Stage Least Squares (2SLS) identifies the Local Average Treatment Effect (LATE), which is defined as the ATE of the sub-group of compliers (Imbens and Angrist, 1994). With continuous instruments, 2SLS identifies a weighted average of different LATEs. A key takeaway from my MTE results is that the LATE is lower than the ATE, the Average Treated on the Treated (ATT), and the Average Treated on the Untreated (ATUT). In other words, this implies that the extrapolation of LATE obtained from 2SLS would generally underestimate the effect of counterfactual policies that target a different sub-population of compliers.

Importantly, the shape of the MTE curve suggests that targeted policies that can effectively promote a faster return to work of mothers that are unlikely to take a short leave can generate substantial returns. The gap in the propensity to take a short leave by maternal education narrows down as the local availability of childcare increases, which suggests that these “hard-to-reach” mothers come from more disadvantaged backgrounds on average. If this finding can be extrapolated to the U.S., it might help explain why the EITC has strong effects on the employment and wages of single mothers (Nichols and Rothstein, 2016; Kuka and Shenhav, 2020).

Recent examples have shown how the MTE framework provides a general framework to give context to results from different IV’s that appear to be inconclusive or even contradictory (Brinch et al., 2017; Kowalski, 2018). In light of the MTE framework and results, it is not surprising that the literature that relies on difference-in-differences or IV estimates to show causal effects of expanding subsidized childcare or extending the duration of leave on maternal outcomes has found inconclusive findings.

The papers that have analyzed the labor market effects of extending the duration of job-protected maternal leave to two or more years have found that these extensions are associated with negative effects on subsequent maternal labor market outcomes (Lalive and Zweimüller, 2009; Lequien, 2012; Schönberg and Ludsteck, 2014). The extrapolation of my results about the shape of the MTE to these other contexts suggest that the negative intent-to-treat effects are driven by both the selection of mothers that opt to take the longer leaves under the new scheme (which would correspond to mothers with a high unobserved resistance to treatment) and the negative causal effects of longer leaves.\footnote{However, the analysis of leave extensions is complicated by the fact that extending the maximum duration of leave appears to have significant general equilibrium effects (e.g. Thomas (2020) or Xiao (2021)).}

The MTE results can also rationalize why previous papers in the literature have found that the effects of expanding childcare are more important for mothers from disadvantaged backgrounds and single mothers (e.g., Cascio (2009)). Additionally, they provide an intuitive explanation of why initial childcare expansions that start at a very low level of provision (e.g., Baker et al. (2008); Bastian (2020)) have on average bigger effects than subsequent expansions.

**Structure:** I present a brief summary of the German parental leave and childcare systems in section 2. I introduce the data in section 3. I explain the empirical design that I use to analyze how the
duration of parental leave reacts to changes in the local availability of subsidized childcare in section 4. I show the main results from the generalized difference-in-differences estimation in section 5. I then switch gears and introduce the MTE framework in section 6. I present the details on how I implemented the MTE framework in section 7. I show the results of the MTE approach in section 8. Finally, I include some concluding remarks in section 9.

2 Context: Child care and parental leave system in Germany

In this section, I first present the most important characteristics about the childcare system in Germany. I then summarize the expansion process in the availability of childcare between 2007 and 2014. Finally, I explain the main features of the entitlement to a job-protected leave and the parental allowance system in Germany during the period of study.

2.1 Overview of the main characteristics of the childcare system

The childcare system in Germany is highly decentralized compared to other countries (Herzog and Klein, 2018). The federal government sets up minimum education requirements, but each state has its own set of laws regarding the financing, operation, and management of childcare centers. Daycare centers in Germany are almost entirely run by public or not-for-profit private providers, including charities associated to different religious congregations. Both public and not-for-profit private providers receive substantial operating subsidies from the government. On average, parental fees only amounted to around 14%–18% of the total operating costs associated with the provision of childcare for children in the age group 0–3 (Ländermonitor, 2021).

Data about parental fees for the age group 0–3 is in general not available in Germany. Past articles have estimated that the average household out-of-pocket expenditures on childcare fees was around €150 – €250. It represents between 5 to 10% of monthly household earnings (Felfe and Lalive, 2018a; Geyer et al., 2015; Jessen et al., 2020). However, without accurate data it is hard to assess how parental fees changed in the period of study. Furthermore, there is substantial variation in prices across locations, since local authorities are responsible for determining fee schedules. These fees usually vary by household income, number of children, and hours in childcare.

The application and admission process was mostly decentralized during the period of study. In most locations parents had to apply personally to childcare centers. Additionally, not-for-profit childcare centers are generally allowed to define their own admission procedures, so long as they

---

5Only 1.2% of the daycare centers were private-for-profit entities in 2007. This percentage increased to 2.8% in 2014.

6Many states have reduced parental fees in recent years. However, only two states changed the fee system for children ages 0–3 between 2007 and 2016. Rheinland-Pfalz phased in free daycare for children ages 2 and above between 2007 and 2010, while Hamburg abolished fees for children ages 2 and above in 2014 (Busse and Gathmann, 2018).

7The application process has started to become more centralized, especially after the introduction of a legal claim to a seat in childcare in August of 2013. For example, the city-states of Berlin and Hamburg introduced an electronic voucher system in which the households first register for a voucher and then apply to childcare centers on-line for the academic year 2014/2015.
comply with state regulations. This has led to complaints about the lack of transparency in the allocation of childcare seats (Herzog and Klein, 2018).

Finally, I emphasize that given that the private-for-profit sector is negligible in Germany, changes in the local availability of subsidized childcare correspond almost one-to-one to changes in the total availability of childcare in the locality. An analysis of the effects of expanding subsidized childcare in countries like the U.S., where the private-for-profit sector represents a considerable share of the childcare market, would also need to consider the potential responses of the private sector to public investments in childcare (e.g., Berlinski et al., 2020).

2.2 Expansion in subsidized childcare for infants and toddlers

Daycare centers in Germany are classified into four different categories: i) Nurseries (Kinderkrippen), which exclusively take care of children that are between 0 to 3 years old; ii) Kindergärten, which generally admit children that have turned three years old at the beginning of the academic year; iii) After-school daycare centers (Horten) for children that are already enrolled in primary school; and, iv) All-age daycare centers (Kindertagesstätten or Kitas) which admit children from all age-groups.

Between 1992 and 2002, there was a rapid expansion in the number of Kindergärten in Germany (Cornelissen et al., 2018). However, the availability of seats in childcare for children ages 0–3 was still severely limited by 2005, especially in West Germany. In response to the excess demand in seats for infants and toddlers, the federal government introduced the Tagesbetreuungsausbaugesetz (TAG) in 2005. The TAG imposed minimum quality criteria for childcare centers nationwide and pledged to the creation of 230,000 additional seats for children aged 0–3 by 2010. This pledge was strengthened in 2007 in a summit where all the three levels of government committed to expand childcare seats for children ages 0–3 and to reach a participation rate of 35 % nationwide by 2013. Finally, the federal government enacted the Kinderförderungsgesetz (Kifög) in 2008 which established a legal claim to a slot in childcare for all one-year-olds starting from August of 2013.

As part of the Kifög, the federal government allocated up to €2.15 billion to states to partially cover the investment costs associated with the expansion of seats for children ages 0–3. The funds were allocated to states in proportion to the number of children ages 0–3 in 2007 and distributed almost evenly across the period 2008–2013. Although the federal government contributed to the expansion with funding, state and local governments were the actual authorities in charge of the implementation. State authorities were responsible for coming up with the construction and investment guidelines, as well as distributing the funds across counties. Local authorities were responsible for generating demand projections, disbursing the money as investment grants, and supervising the constructions and renovations. Given the administrative and bureaucratic burdens associated with the implementation of such an ambitious project, the timing in the roll-out of Kifög varied substantially across localities.

I present different measures that capture how the availability of childcare for infants and toddlers was expanded in the following subsections. I construct these measures combining pop-

---

8This amount corresponded to an investment of slightly more than €1,000 per child.
ulation counts with information from annual statistical reports that show the status of the child care system by the month of March for each of the 401 counties in Germany.\footnote{Prior to 2007, the statistical reports were only available every four years.}

2.2.1 Number of nurseries and Kitas, per child aged 0–3

Nurseries and Kitas are the only two types of daycare centers that admit infants and toddlers that are younger than two years old. As a first measure of local availability of childcare, I use the number of nurseries and Kitas in the county, normalized by the local number of children in the age group 0–3 (in thousands).

Panel A of Figure 1 shows whisker plots of the normalized number of nurseries and Kitas in the period 2007–2015, separately for East and West Germany. Panel A of Figure 1 shows that the availability of childcare was higher in East Germany than in West Germany in 2007. Additionally, most of the new construction or adaptation of childcare facilities occurred in West Germany. Overall, the average number of nurseries and Kitas per 1,000 children aged 0–3 across counties in West Germany (East Germany) increased from 7.41 (22.15) in 2007 to 14.79 (23.87) in 2015.\footnote{In raw terms, there was a total of 12,559 (7,476) nurseries and Kitas in West (East) Germany in 2007. By 2013, that number had increased to 19,933 (8,356). This change represents a total increase of 58.7\% (11.7\%) between 2007 and 2015.}

2.2.2 Number of childcare seats, per children ages 0–3

A different way to assess childcare availability is with the number of seats per child for children ages 0–3. Unfortunately, the statistical reports do not disaggregate the total number of seats in each county by age group. Using the total number of seats in the locality would provide an inaccurate measure in the availability of childcare for the age group 0–3, as many seats that were originally assigned to older children before Kifög were transformed into seats for infants and toddlers.

Instead, I proceed by obtaining an upper bound on the number of local seats for children in the age group 0–3 by substracting the total number of children in childcare that are older than 3 years old from the total number of seats in the county.\footnote{This measure provides an upper bound in the seats for children ages 0–3 since it relies on the assumption that all the seats that are not taken by children older than three years old can be taken by children in the age group 0–3.} Afterwards, I normalize this measure by the number of children ages 0–3 in the county.

In panel B of Figure 1 I show the evolution of the upper bound of the number of seats per child for the age group 0–3 between 2007 and 2015, separately for East and West Germany. Panel B shows that this variable grew in both West and East Germany during the period of study. In West (East) Germany, the average number of seats per child across counties passed from 0.23 (0.73) in 2007 to 0.47 (0.85) in 2015. Reconciling the evidence from Panels A and B, the expansion in the number of seats for children ages 0–3 occurred within already constructed Kitas in East Germany. Meanwhile, local authorities in West Germany had to construct new Kitas or readapt traditional Kindergärten into all-age group daycare centers. Finally, in terms of the availability of seats there is some overlap between counties in East and West Germany.
2.2.3 Proportion of children in childcare, ages 0–3

The expansion in the availability of childcare was accompanied with a rise in the proportion of children ages 0 to 3 in childcare. The participation rate increased substantially nationwide, from 15.5% in 2005 to 32.3% in 2013. Although the rate more than doubled, the goal of the Kifö to reach a participation rate of 35% by August of 2013 was not reached. Furthermore, the participation rate stagnated in 2014 and 2015, after the flow of federal funds stopped. I show how the participation rates grew across counties in Panel C of Figure 1. The increase in the average participation rate across counties was higher in West (18.1 p.p. increase) than in East Germany (13.0 p.p increase).

2.2.4 Number of pedagogical staff per children, ages 0–7

Finally, I show how the number of pedagogical staff increased between 2007 and 2015. Statistical reports do not disaggregate pedagogical staff by the classroom age group. Instead, I normalize the number of pedagogical staff by the number of children ages 0–7 in the county (in thousands). Given that the childcare participation rate in the age group 3–7 was already above 90 percent by 2007, most of the growth in this variable comes from the growth in the staff assigned to classroom with children ages 0–3. Panel D of Figure 1 shows the evolution of this variable. The average number of pedagogical staff per child in the age group 0 to 7 across counties increased from 64.24 (93.64) instructors per 1,000 children in 2007 to 103.48 (118.95) instructors in 2015 in West (East) Germany.

2.3 Entitlement to a job-protected parental leave

Each employed parent in Germany is entitled to take up to three years of unpaid, job-protected leave at any point between the day of birth of the kid and the day that the kid turns eight years old. In practice, however, most parents take a single period of leave that ends on or before the child turns three years old.

Explicit permission of the employer is only needed if the period of leave starts after the child turns three years old. Nonetheless, employees are required to notify their employers about their leave intentions at least seven weeks before the start of their leave. Employees are protected from dismissal starting from the eighth week before the start of the leave. Hence, most parents in Germany formally notify their employers about their leave plans seven weeks before the end of this mandatory period. This period corresponds to the week after the child is born. For more details, see BMFSFJ.
2.4 Parental allowance

Contrary to the individual entitlement to a job-protected leave, the entitlement to parental allowance is determined at the household level. During the period of study, eligibility for parental allowance was universal. Each couple received a maximum of 14 months of benefits to jointly allocate between the two parents.\textsuperscript{14} However, each parent could only claim a maximum of 12 months of parental allowance on their own.\textsuperscript{15} Finally, the allowance for each parent could only be received within the first 14 months after the child was born.

During the period of study, the parental allowance was calculated from the average monthly earnings of the parent in the year prior to the birth of the child using a replacement rate of 67 percent. The monthly parental allowance was capped at €1,200. A parent with zero or very low average monthly earnings in the year prior to the birth of the child was eligible to receive a monthly parental allowance of €300. Finally, a parent that was receiving parental allowance could also work for a maximum of 30 hours per week. However, any earnings that the parent perceived while receiving parental allowance were deducted from the monthly amount of benefits. This mechanism acted as an important disincentive to work part-time while receiving parental allowance.\textsuperscript{16}

3 Data

3.1 German Child Care Study (KiBS)

I use the German Child Care Study (KiBS) of the German Youth Institute (DJI) to study the interaction between take-up of long parental leaves and childcare availability. The KiBS is a big, representative dataset of households with young children in Germany. Overall response rates hover around 30% across survey rounds. It has been collected since 2012 to analyze the demand for different childcare arrangements in Germany (Alt et al., 2018a,b).

Accurate data on parental leave durations is hard to get. A key advantage of the KiBS dataset is that it includes retrospective information about the duration of parental leave for both parents in the household.\textsuperscript{17} The German Socio-Economic Panel (SOEP) and the National Educational Panel Studies (NEPS) also have information on parental leaves. Nonetheless, the number of households with small children in the KiBS is an order of magnitude higher than in the SOEP or NEPS.

\textsuperscript{14}Two months of parental allowance are earmarked for fathers since 2007. Take-up of paternal leave has risen gradually since the introduction of these two “daddy months”, similar to what happened in Norway (Dahl et al., 2014). The maximum number of months of parental allowance increases automatically to 14 months for single parents.

\textsuperscript{15}Intact households lose two months of benefits if fathers do not reduce their working hours. The literature informally refers to this mechanism to incentivize paternal leave by earmarking benefits for fathers as “daddy months”.

\textsuperscript{16}The parental allowance was reformed in January of 2015 to allow for more flexibility in the time and way in which parents are able to receive parental benefits. Under the new system, parents can choose between receiving full benefits for a given month or distributing these benefits into two months where there is no disincentive to work part-time.

\textsuperscript{17}The Microcensus only asks questions about current employment. Similarly, identifying parental leave spells using administrative datasets has significant challenges. Although some of the out-of-work spells related to childbirth can be inferred for mothers, these spells cannot be further classified into spells where the mother was on leave and spells where the mother was out of the labor force. In the case of men, the duration of paternal leave is only observed for fathers that took some leave. Couples identifiers are not included, which makes the analysis of joint leave behavior impossible. Finally, the actual number of hours worked is not available in the administrative datasets.
In the first four years of data collection, the sampling design of the KiBS dataset was limited to households with children under the age of three. Around 800 households were surveyed in each state per year. In 2016 and 2017 the sampled population was expanded to include close to 24,000 households with children ages 3–14. I focus on the data from the waves in 2016 and 2017 to obtain the duration of parental leave spells when the focal child is at least three years old.

The public version of the KiBS only has state identifiers. For this project the DJI shared two restricted variables that help characterize the location of households more precisely: the type of county and region by urbanicity level, and an indicator variable on whether the municipality of residence has a population of 500,000 inhabitants or more. In all what follows, localities are defined as the intersection between state ID’s × urbanicity of county × urbanicity of region × indicator on whether the municipality has 500,000 or more inhabitants.

Based on this definition, counties are grouped into 89 different locations. To match the definition of the different locations across datasets, I also aggregate all the relevant childcare statistics from Section 2 to this new level of aggregation. Figure 2 shows a map of all 401 counties by type. Unfortunately, this means that the clustering of different counties into a same location introduces measurement error in the availability of childcare.18

3.2 Sample selection

I define year cohorts based on the German school year, which starts between July and September in almost all states. I pool households where the child was born between August of year \( y - 1 \) and July of the following year \( y \) into the same year cohort \( y \). To avoid the major reforms to the parental leave system in January of 2007 and 2015, I restrict the analysis to households where the focal child belongs to the 2008 - 2014 birth cohorts, \( i.e. \), to households where the focal children were born between August of 2007 and July of 2014.

The KiBS dataset does not contain the full employment history of the parents before the birth of the child. However, I proxy for the eligibility to a job-protected maternity leave using the reported duration of maternal leave. As was mentioned in Section 2, all working mothers are mandated to take 8 weeks of maternity leave after the birth of the child. Hence, I identify eligibility for a job-protected maternal leave if the reported maternity leave is positive.19 In most parts of the analysis, I restrict the sample to the subset of households where mothers are eligible to take a job-protected leave based on this criterion. I further limit the sample to households where the mothers were at least 22 years old at the time of birth to reduce the probability that mothers were in school or in an apprenticeship at the time of birth. Finally, I drop some households where basic

---

18This measurement error is not even across states. In the most populated states of North Rhine-Westfalia, Bavaria, and Baden-Württemberg, there are more counties aggregated into a single location than in the least populated states of Saarland and Bremen.

19Depending on how they understood the survey question, some mothers might have opted to report their leave duration without considering the first 8 weeks of mandated leave. Hence, it is possible that some mothers that did not take any additional leave apart from the 8 weeks of mandated leave are misclassified into not being eligible for parental leave. However, this misclassification error is most likely negligible. Less than 1% of the mothers reported receiving only 1 or 2 months of parental allowance nationwide.
demographic information is missing.

The main sample of households where mothers were eligible to take a job-protected leave has 12,643 observations. In column (1) of Table 1, I present some summary statistics of the main sample. The average maternal leave duration is equal to 17.14 months. Meanwhile, 10.7% of the mothers are foreign-born. Following Cornelissen et al. (2018), I distinguish between foreign-born mothers that were born in Eastern Europe and Russia and foreign-born mothers that were born in the rest of the world. I measure the education level of mothers based on whether they passed the Abitur, which is a type of secondary education degree that enables students to attend university. 58.7% of mothers in the main sample passed the Abitur. Finally, the average age of mothers at the time of birth of the focal child is 32.74 years.

In Table 1 I also present summary statistics related to maternal labor market outcomes which are measured at the time of the interview, when the focal child is 5 years old on average. Almost 75% of the mothers report working 10 or more hours per week. The average maternal weekly hours of work in the sample is 22.46, and 6.4% of the households report receiving transfer payments. Finally, the average maternal monthly income equals €1,216.6, including zeroes. Without considering the mothers that report working less than 10 hours per week, the average monthly income of mothers rises to €1,606.1.

In Figure 3 I show how the cumulative distribution function of the maternal leave duration evolved from the 2008 birth cohort to the 2014 birth cohort. Panel A shows the evolution of the duration of maternal leave for eligible mothers in West Germany, while Panel B shows it for eligible mothers in East Germany. From the two graphs, it is clear that the cumulative distribution function of the duration of maternity leave has shifted to the left over time.

However, the proportion of mothers that took less than 12 months of leave only changed slightly over the period of study nationwide (14.2% for mothers in the 2008–2010 birth cohorts, compared to 14.9% in the 2013–2015 birth cohorts). This provides suggestive evidence that there was no significant change in the proportion of mothers that forewent some of the 12 months of parental allowance at the national level. In contrast, there was a considerable increase in the proportion of mothers that took exactly 12 to 14 months of leave (45.2% in the 2008-2010 birth cohorts, compared to 52.1% in the 2013–2015 birth cohorts). From the comparisons of Panel A and Panel B, this increase mostly came from mothers in West Germany. This indicates that many mothers in West Germany shifted from taking 24 or 36 months of leave to taking 12 months of leave during the period of study.

3.3 “Short” and “long” leaves

Based on the evidence from Figure 3 and the strict window of 14 months after the birth of the child that parents had to claim federal parental allowance in the period of study, I also classify leaves in a dichotomous fashion. In what follows, I classify leaves as “short” if the total duration of the

---

20 I define that a household is receiving transfer payments at the time of the interview if the respondent reports that any member of the household is receiving any of the following social benefits: either short-term or long-term unemployment, social assistance (Sozialhilfe), or housing benefits (Wohngeld).
maternal leave is at most 14 months, and “long”, otherwise. Short leaves are most likely fully paid. In contrast, mothers that took long leaves most surely did not receive federal parental allowance for at least some of the months on leave.

The second and third columns of Table 1 present summary statistics of households based on the type of leave of the mother. The comparison between these two columns show important differences in the average demographic characteristics and subsequent labor market outcomes of the two types of households. Mothers that went on short leaves are on average younger and more educated, slightly less likely to be foreign-born, and more likely to live in locations that have a higher number of nurseries and Kitas per child. In terms of labor market outcomes at the time of the interview, mothers that took short leaves on average work more hours per week, earn more per month, and are more likely to be working at least 10 hours per week.

In Sections 4 and 5, I will assess whether the association between the duration of maternal leave and the local availability of childcare that is reported in Table 1 is causal. Instead, in Sections 6 and 7 I will quantify how much of the observed differences in maternal labor market outcomes at the time of the KiBS interview between mothers that took short leaves and mothers that took long leaves is driven by the causal effects of short leaves on outcomes and the selection of mothers with different observed and unobserved characteristics to different maternal leave durations.

4 Effects of childcare availability on leave outcomes: Empirical strategy

4.1 Generalized difference-in-differences design

I present the empirical design that I use to analyze the effects of expanding subsidized childcare on parental leave outcomes in this section. A commonly used approach in the literature to analyze the effects of expanding childcare for children of a given age group on family outcomes is to use a Difference-in-Differences (DiD) specification. In DiD, locations are clustered into two different groups (“treated” or “untreated”) based on the introduction or expansion of childcare programs across different states (Baker et al., 2008; Cascio, 2009), or on the relative speed or intensity in the implementation of a federal expansion across localities (Havnes and Mogstad, 2011; Bauernschuster and Schlotter, 2015).

In the case of Germany, the parental allowance system was completely reformed in January of 2007. This change matches up with the time in which the expansion in childcare seats for infants and toddlers was ramping up. This means that the parental allowance system in place in the period of time before the childcare expansion (pre-treatment period) is completely different to the parental allowance system in the period of time after the start of the childcare expansion (post-treatment period). The interpretation of the results of a DiD specification might conflate the effects of expanding childcare availability with the effects of modifying the parental allowance system, even after controlling for time effects.

Instead of pursuing a DiD specification, I analyze how changes in the local availability of childcare for children in the age group 0–3 affect household decisions using a generalized Difference-
in-Differences design that follows the specification in Berlinski and Galiani (2007) and Müller and Wrohlich (2020). More specifically, I assume that there is a linear relationship between the local availability of childcare for children ages 0–3 represented by $Z_{liy}$ and household outcome $Y_{ilyq}$ as follows:

$$Y_{ilyq} = \beta Z_{liy} + \alpha X_{ilyq} + \alpha R_{ly} + \gamma l + \gamma yq + \epsilon_{ilyq}$$ (1)

where subscript $i$ indexes individuals; $l$, the location of residence at the time of the interview; $y$, the birth cohort year of the focal child; and $q$, the quarter of birth of the focal child. Importantly, the main specification includes birth cohort fixed effects, $\gamma yq$, which are defined by the interaction between cohort year and quarter of birth, and location fixed effects, $\gamma l$. The birth cohort fixed effects absorb national trends in the availability of childcare and household outcomes. Meanwhile, the location fixed effects control for time-invariant differences in childcare availability and household outcomes across localities.

The main parameter of interest is $\beta Z$, which measures how changes in $Z_{liy}$ are associated with changes in $Y_{ilyq}$. The main outcome of interest $Y_{ilyq}$ is the total duration of maternal leave. However, I also analyze the effect of changes in the availability of childcare on other outcomes, including the duration of paternal leave. I also classify leaves into short and long leaves, as has been explained in Subsection 3.3. For this discretized outcome, I assume that the relationship between the probability of taking a short leave rather than a long leave and the local availability of childcare follows a probit specification:

$$\Pr (Y_{ilyq} = 1) = \beta Z_{liy} + \alpha X_{ilyq} + \alpha R_{ly} + \gamma l + \gamma yq + \epsilon_{ilyq}$$ (2)

where $\epsilon_{ilyq} \sim N(0, \sigma_\epsilon)$.$^{21}$

Both regressions (1) and (2) control for a vector of observed household characteristics $X_{ilyq}$. In the main specification, $X_{ilyq}$ includes an indicator variable on whether the mother passed the Abitur, the age of the mother at the time of birth, and two indicator variables on whether the mother is foreign-born, distinguishing between mothers born in countries from the former Soviet Union and Turkey, and all other foreign-born mothers.$^{22}$ The specifications also allow for the inclusion of time-varying characteristics at the local level, $R_{ly}$, that can affect household leave decisions. Changes in the economic prosperity or the overall fertility behavior in the locality can influence household leave decisions. Thus, in some specifications I also control for the local unemployment rate and the ratio of women ages 15–45 that gave birth during the year in the locality.

It is also possible that the expansion in childcare access was accompanied by changes in the quality of the childcare services provided. In order to control for this, I construct an index of quality at the local level using the total number of children in childcare by age groups, the total number of pedagogical staff, and optimal student-teacher ratios. This quality index is defined as the ratio of

$^{21}$I also estimate linear probability models and obtain virtually identical results. The probit specification has the advantage that the predicted propensity score for each observation is between 0 and 1. This trait eases the estimation of the Marginal Treatment Effects in Section 6.

$^{22}$Cornelissen et al. (2018) show that parents from the former Soviet Union and Turkey were less likely to send their children to kindergarten than other minorities between 1994 and 2006.
observed pedagogical staff to the pedagogical staff that would be required to fulfill predetermined optimal student-teacher ratios. I control for this quality index in some specifications.

### 4.1.1 Measure of local availability of childcare

Previous related literature in Germany has used the proportion of children that is attending publicly subsidized childcare in a cohort as a proxy for the number of available seats per child (Bauernschuster et al., 2016; Müller and Wrohlich, 2020). If the seats in some localities are not rationed, then it is possible that changes in the proportion of children in childcare might be driven by changes in demand instead of supply. This would raise an endogeneity concern.

Thus, I use the number of nurseries and Kitas per 1,000 children in the age group 0–3 in the locality as my preferred measure of local availability of childcare. To construct this measure, I aggregate the number of nurseries and Kitas in all the counties that belong to a given location, and divide it by the aggregate number of children in the locality. This variable is slow-moving, so it is unlikely to react to sudden changes in demand. Additionally, the construction of new daycare centers or the transformation of traditional Kindergärten into Kitas is determined by many different local factors, including administrative red tape, differences in land costs, or difficulties in securing construction lots, that are plausibly uncorrelated to the excess demand of childcare.

Another alternative would be to use the number of seats per child for the age group 0–3, like Andresen and Havnes (2019) for the childcare expansion in Norway. However, as I explained in subsection 2.2.2, the statistical reports do not disaggregate the number of seats in the county by the age group of the child. However, I use the upper bound of the number of childcare seats for children ages 0–3 that was presented in subsection 2.2.2 as an alternative measure of availability in unreported validation checks.

Needless to say, the number of different childcare centers per child and the number of seats per child in the locality capture different aspects of childcare availability. For example, the number of nurseries and Kitas is potentially more informative about the average distance of a randomly selected household to the nearest childcare center than the number of seats per child. Additionally, childcare centers vary widely by the type of services that they provide, their management style, or their approach (e.g., bilingual or monolingual, Waldorf or Montessori, affiliated to a religious organization or secular, part-time or full-time). The number of nurseries and Kitas is also more informative about this heterogeneity in services. Holding fixed the number of seats per child in the locality, both of these aspects can affect the probability that parents in a given household find a childcare option that suits their preferences.

I consider that the optimal student-teacher ratios are 3:1 for children ages 0–3, 8:1 for children ages 3–7, and 12:1 for children in after-school care. These ratios are quite close to the recommended ratios of organizations like the National Association for the Education of Young Children (NAEYC) in the U.S. and the Bertelsmann Stiftung in Germany. The exact formula that I use to construct the quality index in a given locality \( l \) in year \( y \) is the following:

\[
\text{Index}_{l_y} = \frac{\text{Pedagogical staff}_{l_y} \cdot \text{Children in childcare ages 0–3}_{l_y}}{3} + \frac{\text{Children in childcare ages 3–7}_{l_y}}{8} + \frac{\text{Children in childcare ages 7–14}_{l_y}}{12}
\]
4.1.2 Timing

As mentioned in Section 2, statistics about the childcare system at the local level are only reported once per year. These statistics correspond to the status of childcare by the month of March. Furthermore, most mothers inform their employers about their leave plans during the first week after the child is born (for details, see subsection 2.3). Finally, I have defined previously that the children born in cohort year $y$ correspond to all the children born between August of year $y - 1$ and July of year $y$.

I proceed by assigning the measure of childcare availability in March of year $y$, $Z_{l,y}$, to the households with children that belong to cohort year $y$. This assignment is the one that most closely assigns households to the local availability of childcare at the time when the relevant parental decision has to be made. This implicitly assumes that parents are myopic and that there is perfect commitment in leave durations. A similar timing convention is used for the other time-varying characteristics of localities, $R_{l,y}$. Nonetheless, I also report additional robustness checks in which I add lagged or future values of $Z_{l,y}$ to the main specification and verify that these inclusions do not substantially change the regression results.

5 Effects of childcare availability on leave outcomes: Results

5.1 Graphical approach

The source of exogenous variation that identifies parameter $\beta_Z$ of equation (1) stems from changes in the normalized number of nurseries and Kitas, after netting out the effect of national trends, permanent differences across locations, and observable household characteristics. Before presenting the results of the linear generalized difference-in-differences specification, I first present binned scatterplots of the residualized versions of the maternal leave duration and the local availability of childcare in Figure 4. These binned scatterplots help visualize the underlying relationship between the two variables, without the imposition of further parametric assumptions.

Residuals in Panel A of Figure 4 come from regressions that only include birth cohort and locality fixed effects. Residuals in Panel B of Figure 4 are instead obtained from regressions that also control for observable household characteristics. In both graphs, all the observations are ranked and grouped into 20 different bins, based on the residualized measure of childcare availability. On the y-axis, Panels A and B of Figure 4 show the average of the residualized maternal leave duration for each bin. The graphs also include a histogram of the residualized measure of local availability of childcare to show the underlying variation in the data. For graphical purposes, I use a 98% winsorization of the residuals of childcare availability.

Figure 4 provides evidence that there is a negative relationship between childcare availability and the duration of maternal leave. Furthermore, Figure 4 indicates that this relationship is not driven by outliers. From the comparison between the graphs in Panel A and Panel B, the relationship appears to be virtually unchanged after controlling for observable household characteristics. Finally, the underlying variation in the availability of childcare is substantial. The 25th and 75th
percentiles of the distribution of the residuals of childcare availability correspond to -0.51 and 0.36 Kitas per child in the age group 0–3.24

In Appendix Figure A-1 I show the same binned scatterplots for the three alternative measures of the local availability of subsized childcare: the upper bound on the local number of seats per child for children in the age group 0–3, the total number of pedagogical staff per child aged 0–7 in the locality, and the proportion of children ages 0–3 that attends subsidized childcare. In all three cases, the relationship between the residualized maternal leave duration and the alternative measure of availability of childcare is negative and not driven by outliers. This provides support to the claim that the negative relationship between the local availability of subsidized childcare and maternal leave duration is not driven by the choice of availability measure.

5.2 Estimates of the generalized difference-in-differences

Table 2 presents the results of estimating equation (1) using OLS. The specification in column (1) only includes the number of local nurseries and Kitas per 1,000 children in the age group 0–3, apart from locality and birth cohort fixed effects. The specification presented in column (2) includes observed household characteristics. Estimates in column (3) come from a specification that additionally controls for the local unemployment rate and the number of births per women in the age group 15–45 in the locality. Finally, estimates in column (4) control for the local quality of childcare using the quality index explained in Section 4. Since the exogenous source of variation occurs at the locality level, I cluster the standard errors at the locality level in all the specifications, following Bertrand et al. (2004).

My preferred specification is presented in column (2). Based on the results of Table 2, an additional nursery or Kita per 1,000 children in the age group 0–3 reduces the maternal leave duration by 0.43 months (13 days) on average.25 This coefficient is statistically significant at conventional levels. The mean leave duration in the sample is equal to 17.1 months. Thus, one additional childcare center per 1,000 children in the age group 0–3 causes a decrease of 2.4% of the overall mean leave duration in the period of analysis.

The comparison between columns (1) and (2) shows clear evidence that the inclusion of observable household characteristics does not affect the estimates of $\beta_Z$. This result provides suggestive evidence that changes in the local availability of childcare supply are orthogonal to observable household characteristics. Finally, the results in columns (3) and (4) show that the inclusion of other time-varying characteristics at the local level does not substantially affect the estimate of $\beta_Z$. Furthermore, I fail to reject that the coefficients from these local characteristics are jointly equal to zero at conventional significance levels.

Even though observable household characteristics appear to be uncorrelated with changes in the local availability of childcare for infants and toddlers, the estimates from column (2) show

---

24 The 25th and 75th percentiles of the measure of seats per child correspond to -0.015 and 0.015 seats per child.
25 The average number of seats in a nursery or Kita is 74 seats nationwide. The average remained stable over time. However, only around 20-25% of the seats in nurseries and Kitas are assigned to children 0–3. Hence, an additional nursery or Kita per 1,000 children represents an increase of approximately 0.017 additional seats per child.
that these characteristics have nonetheless an important effect on the duration of maternal leave. Mothers with an abitur certificate and younger mothers take shorter leaves on average.

5.3 Discretized maternal leave duration

In order to analyze whether the availability of childcare affects the propensity of mothers to go on “short” leaves, I estimate different versions of equation (2) which assumes a probit structure. The new dependent variable is an indicator variable on whether the mother took a leave that was 14 months or shorter. Table 3 presents the estimated average marginal effects of this probit specification. I follow the same structure and included covariates of Table 2. Column (1) presents the estimates of a specification that only includes the measure of local availability of childcare, as well as locality and birth cohort fixed effects. Column (2) additionally includes observed household characteristics. Columns (3) and (4) include time-varying characteristics at the local level.

My preferred specification is again shown in column (2). An increase of one additional nursery or Kita in the locality (per 1,000 children in the age group 0–3) raises the probability that mothers go on a short leave by 1.6 percentage points on average. Again, younger mothers and mothers with an abitur certificate are more likely to choose a short leave. In contrast, mothers that were born in Eastern Europe are less likely to take short leaves. The comparison between column (1) and column (2) again shows that the inclusion of observed household variables does not affect the estimate of $\beta_Z$.

Column (3) shows that the estimated coefficients on the local unemployment rate are now negative and statistically significant at conventional levels. In particular, an increase in the local unemployment rate by 1 percentage point is associated with a decrease in the probability that mothers take a short leave of 2.3 percentage points. This result is consistent with a story in which a worsening in the local market conditions is associated with a decrease in the economic benefit of taking a short leave. Finally, in column (5) I present the results of the baseline specification using a linear probability model (LPM). The estimated average marginal effect from the LPM and the probit models are virtually identical.

In Appendix Figure A-2 I show how the estimated average marginal effect and its 95 percent confidence interval vary as I modify the cut-off to classify short and long leaves from a minimum cut-off of 6 months to a maximum cut-off of 35 months. Appendix Figure A-2 shows that the estimated marginal effect reaches its peak at the 12th and 13th months, quite close to the conservative cut-off of 14 months that I use. This result is consistent with a story in which many mothers shifted from taking a maternal leave of 24 or 36 months to taking a 12-month maternal leave.

5.4 Exploring the effects of increasing childcare availability on other outcomes

I also analyze the effects of expanding access to subsidized childcare on the joint leave decision of both parents. I restrict the analysis in this subsection to the sub-sample of households that remain
intact at the time of the survey.\textsuperscript{26} Close to 92 percent of all focal children in the main sample were living with both biological parents at the time of the survey.

Cohabitation might also react to changes in the local availability of childcare. Thus, I first analyze whether the local availability of childcare affects the probability that children are observed in intact households at the time of the survey. This effect is ex-ante ambiguous. On the one hand, childcare availability might raise the bargaining position of mothers relative to fathers (Doepke and Kindermann, 2019). If the change in bargaining power from an increase in the local availability of childcare is big enough, it might lead to a negative effect on the probability of observing intact households at the time of the interview. On the other hand, childcare availability might mitigate coordination problems arising from the assignment of childcare responsibilities within partners. This other explanation would suggest a positive effect of childcare availability on couple stability.

To analyze the effects of changes of local availability of childcare on the probability that households remain intact, I estimate equation (1) now using as dependent variable an indicator variable that is equal to one if the focal child is living with both biological parents at the time of the survey and zero, otherwise. Furthermore, I control for the regressors included in the main specification of the maternal leave duration. Unlike the measures of leave that have been previously studied, cohabitation status continues to evolve as the focal child ages beyond the age of three. Since there is some variation in the age of focal child at the time of the interview, I add the age of the child at the time of the survey as an additional regressor.

The results are shown in column (1) of Table 4. In particular, I find that one additional childcare center per 1,000 children ages 0–3 in the locality reduces the probability that focal children are observed living with both biological parents by 0.60 percentage points. This effect is statistically significant (p-value = 0.04). Two different stories could explain this result. First, the effect of increasing the local level of childcare availability might have a negative effect on household stability after the child is born. However, it is also possible that the availability of subsidized childcare reduces the probability that pregnant mothers cohabit with the biological fathers. Unfortunately, the survey does not contain information about the cohabitation status at the time of birth to disentangle between these two possible alternatives.

In the remaining columns of Table 4, I analyze the effects of expanding childcare centers on leave decisions in the sub-sample of intact households. However, the estimates from columns (2) to (5) do not have a causal interpretation due to the non-random selection of households into cohabitation at the time of the interview. Nonetheless, the extent of selection is most likely small, given that close to 92\% of the children reside in intact households at the time of the interview. In these specifications, I additionally control for the type of school-leaving certificate and nationality of the father.

In column (2) of Table 4, I replicate the main specification shown in column (2) of Table 2 in the sub-sample of intact households. The estimated effect of an additional childcare center per 1,000 children in ages 0–3 (−0.48 months) is slightly bigger in absolute terms than the estimated effect for

\textsuperscript{26}I restrict the analysis to the sub-sample of intact households because information about paternal leave is not asked in the survey if the biological father does not live with the child at the time of the survey.
the whole sample (-0.43 months). Hence, the work and leave decisions of single mothers appear to be less elastic to external factors like the availability of subsidized childcare. However, the sample of single mothers in the KiBS dataset is too small to investigate this further.

In columns (3) and (4) I analyze whether childcare availability has an effect on paternal leave. Column (3) uses as dependent variable an indicator variable on whether the father took any leave. Instead, in column (4) the dependent variable is the number of months that fathers go on leave, including zeroes. The results show that the effect of increasing the availability of subsidized childcare on paternal leave outcomes is positive, albeit economically small and statistically insignificant. For example, I am able to rule out that an additional nursery or Kita per 1,000 children ages 0–3 raises the duration of paternal leave by more than 3 days at the 5 percent significance level. Finally, I verify in column (5) that the expansion of childcare reduces the total number of months of leave at the household level. Hence, I can reject the hypothesis that the reduction in the duration of maternal leave is offset by an increase in the length of paternal leave.

All the results in this sub-section point to the same conclusion. Fathers do not appear to modify their leave decisions in response to changes in the local availability of subsidized childcare. A different result is found for mothers. This conclusion is in line with the literature, which has found that the work and career decisions of mothers are more elastic to exogenous changes in the environment than those of fathers (e.g., Blundell et al. (2016)).

5.5 Heterogeneous treatment effects

The effects of expanding the local availability of childcare might differ based on household characteristics. In this sub-section I estimate heterogeneous treatment effects of expanding childcare on maternal leave duration based on three important household characteristics: region, the level of education of the mother, and the birth order of the focal child.

In Section 2 I described how the conversion of kindergartens to childcare centers for all ages was more prominent in West Germany than in East Germany. It is thus possible that the effects in the main sample are mostly driven by households in West Germany. In order to assess if this is the case, I run the same specifications based on equation (1) using only households in West Germany. Appendix Table A-1 replicates the results of Table 2 using localities in West Germany only. Column (2) of Appendix Table A-1 shows my preferred specification. The estimated effect of changes in local availability of childcare on maternal leave duration in West Germany is quite similar to the overall effect in Germany: an increase in one childcare center per 1,000 children ages 0–3 reduces maternal leave duration by 0.358 months (11 days). Again, this coefficient is stable across columns (1) through (4). In a specification that fully interacts all variables with an indicator of East Germany, I fail to reject the null that the coefficient $\beta_Z$ is the same for the two regions (p-value = 0.360).

Another interesting margin of heterogeneity to study corresponds to the level of education of the mother, based on whether the mother passed the Abitur or not. The effects of increasing childcare availability for mothers with a high level of education might be either higher or lower than
the effects for mothers with a low level of education depending on how the childcare administrators prioritize seats when there is excess demand. It is also possible that the insurance value of a long leave when the availability of childcare is low is higher for mothers with low levels of education.

I present the estimated effects of an additional nursery or Kita (per 1,000 children ages 0–3) in the locality, separately by the level of education of the mother, in Panel A of Table 5. Column (1) replicates the main results for the duration of maternal leave using OLS. Column (2) presents the results of estimating equation (1) in the sample of mothers that passed the Abitur, while column (3) shows the estimates for mothers without the certificate. The results show that the effects on maternal leave duration and propensity to take a short leave are statistically different by type of education. The estimated effect of increasing the local availability of childcare by one unit on maternal leave for mothers with (without) an abitur certificate is -0.630 (-0.227) months. The difference between the groups is statistically significant (p-value = 0.015).

Finally, a key variable that might be crucial in determining parental leave durations is the child’s parity. For example, prior research has found that childcare expansions have bigger effects on the employment outcomes of mothers with older children (Nollenberger and Rodríguez-Planas, 2015). The KiBS dataset does not have information on the number of children at the time of birth of the focal child. However, a substantial proportion of households that were surveyed in 2016 or 2017 had been previously contacted in the first four waves of the KIBS, when the focal child was between 0 and 35 months old. From these 8,987 households that were observed at least twice, I use the number of kids in the household in the first interview as a proxy for the parity of the child.

Panel B of Table 5 shows the heterogeneous effects of changes to the local availability of childcare on the duration of leave, by the number of children in the household at the time of the first interview. I aggregate households into two different groups: households with only the focal child, and households with more children. Column (1) replicates the baseline result in the sub-sample of households that are observed at least once before the child turns three. The estimated effects are similar to the baseline sample. Column (2) shows the estimated effects of expanding childcare at the local level for households that report having at least two kids in the household at the time of the first interview, while Column (3) shows the same effects for households where the focal child is the only child present in the first interview.

The estimates in Panel B of Table 5 provide evidence that childcare expansions have a bigger effect on maternal leave durations in high parity births. The difference between both groups is close to being statistically significant at the 10 percent level (p-value = 0.116). One possible explanation behind this empirical pattern is that first-time mothers are more attached to the labor force than mothers with two or more kids. This is also confirmed by the differences in mean durations of both groups. However, it is hard to tease out specific mechanisms, as mothers that have a second child and that are eligible to take parental leave might differ from first-time mothers in multiple ways.
5.6 Validation and robustness checks

The main assumption underlying the interpretation of $\beta_Z$ as causal is that:

$$\mathbb{E}[\varepsilon_{ilyq} \mid Z_{ly}, X_{ilyq}, R_{ly}, \gamma_1, \gamma_y] = 0 \quad (3)$$

This assumption cannot be tested directly. Nonetheless, in this section I argue why equation (3) most likely holds in my set-up. I first summarize some key empirical findings from previous papers that validate my research design in subsection 5.6.1. In the remaining subsections I provide novel evidence that the variation stemming from the childcare expansion is arguably exogenous.

5.6.1 Prior literature

Recent papers have exploited the same source of plausibly exogenous variation stemming from the differential roll-out of the childcare expansion across locations in Germany to study the effect of childcare availability on different outcomes (Bauernschuster et al., 2016; Felfe and Lalive, 2018a; Müller and Wrohlich, 2020). Many of the arguments that validate the exogeneity of the reform in those papers apply directly to my set-up. Instead of redoing part of their analyses, I opt to briefly summarize their findings in this subsection.

A first major endogeneity concern is that counties that had a faster or more intense childcare expansion were on a different path in terms of their time-varying characteristics than counties with a slower or less intense childcare expansion. However, Müller and Wrohlich (2020) provide empirical evidence that there were no differential pre-trends in the employment status of mothers with children ages 1–3 in counties where the childcare expansion was above median intensity compared to counties with a below median intensity in West Germany. Felfe and Lalive (2018a) extends the analysis of pre-trends to a bigger set of time-varying characteristics. Similar to Müller and Wrohlich (2020), the authors do not find any evidence that the evolution of different local characteristics in counties with a fast implementation was any different to that observed in counties with a slow implementation in the state of Schleswig-Holstein.

Other endogeneity concerns, like selective migration and changes in the composition of mothers, have been discussed in Bauernschuster et al. (2016).27 Bauernschuster et al. (2016) does not find any evidence that local changes in the in-migration rates of women of reproductive age are systematically linked to local changes in the childcare coverage using administrative data on inter-county migration flows from West Germany. Furthermore, Bauernschuster et al. (2016) provides suggestive evidence that it is unlikely that the childcare expansions modified the composition of women giving birth by showing that there is no discernable association between changes in the health outcomes of children at birth and changes in the local childcare coverage.

27Selective migration is an important concern in my set-up, as I only observe the location of a household at the time of the survey and not the location of a household when the child is born.
5.6.2 Time-varying unobserved characteristics at the locality level

Even if counties with fast and slow childcare expansions were on similar trends in terms of observed characteristics, it is still plausible that they were on different trends in terms of unobserved characteristics. For example, it could still be the case that counties with a fast childcare expansion were also implementing other policies at the local level that promoted a faster return to work for mothers. I perform a validation check similar to one used in Johnson and Jackson (2019) in the context of Head Start spending that provides suggestive evidence that this is unlikely.

The effects of childcare availability on maternal leave durations should theoretically only have an effect at the time when mothers are required to inform their employers about their leave plans. This notification happens in the first week after the child is born. After controlling for the childcare availability in this critical window of time, future or lagged values of childcare availability should not influence leave outcomes. However, if localities with a faster childcare expansion had also implemented or were in the process of implementing other policies related to leaves, future or lagged values of childcare availability would capture these unobserved changes at the local level.

Thus, I test whether prior or future levels of childcare availability have an effect on the duration of maternal leave, after conditioning on the childcare availability associated to the relevant year for each child. To do this, I include either one future or lagged value of childcare availability to my preferred specification of equation (1). In Figure 5 I present the point estimates and 95% confidence intervals of the coefficients associated to these lagged or future values. Each estimate comes from a separate regression where only that lead or lag was included.

Figure 5 shows that only the effect of the childcare availability in the year prior to the year of birth is marginally significant (p-value = 0.08), which might be a result of using annual childcare statistics to measure childcare availability. These results provide some extra validation that there are no time-varying unobserved characteristics at the local level confounding the main results.

5.6.3 Unobserved changes in the demand for childcare

Another related endogeneity concern that would violate equation (3) is that the childcare availability measure captured by $Z_{ly}$ is partially reacting to the unobserved demand for childcare instead of exclusively changing in response to supply considerations. This could be an important concern if the measure of childcare availability can react fast enough to changes in demand, like the observed proportion of children in childcare, or even the number of seats per child in the locality. However, my preferred measure of local availability of childcare $Z_{ly}$, the normalized number of nurseries and Kitas in the locality, is slow-moving and unlikely to react immediately to sudden changes in the local demand for childcare.

5.6.4 Selection into the sample

An important issue in the interpretation of the results shown in Section 4 is the presence of differential selection into motherhood. Expanding the availability of childcare might change the
composition of women that select into giving birth. In my set-up, an additional margin of selection
involves the composition of women that selected into giving birth that were also eligible to take a
job-protected leave. Given the eligibility criteria of job-protected leaves in Germany, this margin
of selection corresponds to all the women that were attached to an employer prior to giving birth.

I complement the results of Bauernschuster et al. (2016) that differential selection into mother-
hood does not appear to be an important concern by following the empirical strategy presented in
Cornelissen et al. (2018). If changes in the childcare availability at the local level are uncorrelated
with changes in observed characteristics of households in the sample of mothers or sub-sample of
mothers eligible to parental leave, then it is plausible that changes in the local level of childcare
availability did not change the composition of women that self-selected into motherhood or into
the sample of mothers eligible to take parental leave.

In order to test whether changes in the local availability of childcare and changes in observed
household characteristics are correlated, I run equation \( (1) \) now using the availability of childcare,
\( Z_{i,y} \), as the dependent variable instead of as an independent variable. I run this regression in two
samples: (1) all mothers regardless of parental leave eligibility,\(^{28}\) and (2) the sample of mothers
entitled to parental leave, which was the main sample used throughout Section 4.

The estimates of these regressions are shown in Table 6. Column (1) shows the results for the
sample that includes all mothers, while column (2) shows the results for the sub-set of mothers
that were eligible to take a job-protected leave. The results from these two columns show that
none of the estimated coefficients are statistically significant. Additionally, I fail to reject the null
hypothesis that the coefficients on all the demographic variables in the regression are jointly equal
to zero. The p-value of these tests are equal to 0.600 and 0.745. Overall, these results suggest that
the changes in the composition of mothers and mothers entitled to a job-protected leave are most
likely small.\(^{29}\)

Instead, changes in the composition of mothers that took short leaves relative to long leaves
in response to expanded childcare availability might be expected. For example, in subsection 5.5
I found evidence that the effects of increasing childcare availability on the duration of maternal
leave were bigger for mothers that did not have an Abitur compared to the effects of mothers that
had passed it. In column (3) of Table 6 I limit the same analysis to the selected sub-sample of
mothers that took a short maternal leave duration. In this case, I do see that changes in the local
availability of childcare are negatively correlated with the level of education of the mothers in this
selected sub-sample. In other words, this indicates that the composition of mothers that decide to
take short leaves changes in response to a higher level of childcare availability.\(^{30}\)

\(^{28}\)I continue to limit the analysis to all mothers that were 22 or older at the time of birth.

\(^{29}\)In unreported results, I also run regressions where the dependent variable is a given demographic characteristic
and the controls are \( Z_{i,y-1} \) and locality and birth cohort fixed effects. In these regressions I again find that \( Z_{i,y-1} \) is not
correlated with any of the main demographic characteristics, after controlling for locality and birth cohort fixed effects.

\(^{30}\)I thank Joseph Mullins for suggesting this analysis.
6 Marginal Treatment Effect framework

Two main mechanisms can explain why there are important differences in subsequent labor market outcomes of mothers that took short leaves compared to mothers that took long leaves. The first explanation is selection, where mothers with different observed and unobserved characteristics self-select into different leave durations. Importantly, these traits can also influence subsequent labor market outcomes. The second one is that long leaves, relative to short leaves, have causal effects on subsequent labor market outcomes.

I adopt the Marginal Treatment Effect (MTE) framework first presented by Björklund and Moffitt (1987) and further developed by Heckman and Vytlacil (1999, 2001, 2005, 2007) to analyze the implications of selecting treatment choice $D$ on subsequent labor market outcome $Y$. Motivated by the empirical findings from the previous section, I define the treatment choice $D$ as follows: $D = 1$ if the duration of maternal leave is at most 14 months (“short leave”), and $D = 0$ if the duration of maternal leave is 15 months or longer (“long leave”). I analyze the effects of this treatment choice $D$ on different labor market outcomes $Y$ after the focal child turns three years old, including an indicator variable on work status, the total number of hours worked, the monthly maternal income, and an indicator variable on whether the household received transfer payments.

The labor market outcome $Y$ observed by the econometrician depends on the choice of treatment $D$ as follows:

$$Y = DY_1 + (1 - D) Y_0$$

where $Y_d$ represents the labor market outcome that a household would obtain after selecting treatment choice $D = d$, where $d \in \{0, 1\}$. Following Heckman and Vytlacil (2005), I assume that the potential outcomes equations $Y_0$ and $Y_1$ can be written as an additively separable function in observed ($X$) and unobserved household characteristics ($U_d$):

$$Y_d = \mu_d (X) + U_d, \quad d \in \{0, 1\}$$

that satisfies the conditional mean assumption $E [U_d | X] = 0$.

The MTE framework allows households to be heterogeneous in terms of $U_1$ and $U_0$, which represent the unobserved components of the labor market outcome of interest $Y$ after a short leave or a long leave, respectively. Importantly, these variables are not readily observed by the econometrician. As an example, when analyzing whether mothers are attached to the labor market after the child is old enough to enroll in kindergarten, $U_d$ might represent, among other things, the promotion prospects of the mother after a leave associated to treatment branch $d$.

Based on the previous equation, the causal effect of taking a short leave ($D = 1$) compared to a long one ($D = 0$) on subsequent labor market outcomes is given by:

$$Y_1 - Y_0 = \mu_1 (X) - \mu_0 (X) + (U_1 - U_0)$$

The discretization of the continuous maternal leave duration into two treatment options implicitly assumes that the duration of leave is random after controlling for $D$. 

---

31 The discretization of the continuous maternal leave duration into two treatment options implicitly assumes that the duration of leave is random after controlling for $D$. 

---

24
This equation clarifies that the causal effect of the treatment depends on both observed \((X)\) and unobserved household characteristics \((U_1 - U_0)\). Importantly, the MTE framework allows two households that have the exact same observed characteristics to differ in terms of their unobserved gain to treatment.

The choice of treatment branch \(D\) follows the logic of a generalized Roy model. Households choose the treatment alternative that provides them with the highest level of utility. This utility function depends on the households’ observed and unobserved characteristics. Following the standard MTE literature, I write the treatment selection equation as an index model that depends on observed \((X, Z)\) and unobserved household characteristics \((U_D)\). Importantly, the treatment decision not only depends on the observed characteristics \(X\) that influence potential outcomes, but also depends on a set of variables \(Z\) which are excluded from the potential outcome equations.

Following the literature, I assume that the selection equation is additively separable in observed \((X, Z)\) and unobserved characteristics \((U_D)\). Hence, the treatment selection equation can be written down as follows:

\[
D = 1 \ (D^* \geq 0), \quad \text{where } D^* = \mu_D(X, Z) - U_D
\]  

(5)

Two households with identical observed characteristics \((X, Z)\) are allowed to differ in terms of \(U_D\). A higher value of \(U_D\) reduces the probability that a household chooses \(D = 1\). For interpretation purposes, I follow the literature and rewrite selection equation (5) by applying a specific monotone transformation to the inequality inside the indicator function of \(D\):

\[
\mu_D(X, Z) \geq U_D \iff \frac{F_{U_D}(\mu_D(X, Z))}{F(X)} \geq \frac{F_{U_D}(U_D)}{V}
\]

where \(F_{U_D}\) represents the cumulative distribution function of the unobserved trait \(U_D\). The left-hand side of the new expression, \(F_{U_D}(U_D) = P(X, Z)\) corresponds to the propensity score of selecting treatment choice \(D = 1\), while \(F_{U_D}(U_D) = V\) is the quantile of the distribution of \(U_D\) and hence \(V \sim U[0, 1]\). In the MTE literature the term \(V\) is referred as the “unobserved resistance to treatment”.

The biggest advantage of the MTE framework over other methods that require more structure in the household maximization problem is that the MTE framework does not require any additional assumptions about the joint distribution of unobserved characteristics \(\{U_0, U_1, V\}\). The relationship between these three vectors is unrestricted. For example, in my set-up it is likely that \(U_0\) and \(U_1\) are positively correlated. Unobserved factors that affect labor outcomes might not vary substantially based on the duration of leave. Additionally, the generalized Roy model allows unrestricted flexibility in the relationship between \((U_1 - U_0)\), the unobserved gains from treatment, and \(V\), the unobserved resistance to treatment.\(^{32}\)

\(^{32}\)In contrast, most dynamic discrete choice models of female labor force participation implicitly assume that \(V\) and \((U_1 - U_0)\) are uncorrelated when looking at wage outcomes. This is the case because in most models the wage differences between short and long leaves would only come from different levels of accumulated experience (e.g. Adda et al. (2017)).
The key data limitation that permeates through the treatment effects literature is that individuals are only observed in one treatment branch. It is impossible to estimate individual causal effects of the treatment. Instead, the treatment effects literature has focused on different conditional expectations of the causal effects that are meaningful for econometricians and/or policy-makers. In particular, the MTE function has been defined by Heckman and Vytlacil (1999) as:

\[
\text{MTE}(x, v) = E[Y_1 - Y_0 | X = x, V = v] = \mu_1(x) - \mu_0(x) + E[U_1 - U_0 | X = x, V = v]
\]

Thus, the \( \text{MTE}(x, v) \) is defined as the average treatment effect of households with observable characteristics \( x \) and unobserved resistance to treatment \( v \). Heckman and Vytlacil (1999, 2005, 2007) have shown that the MTE is an incredibly useful concept to identify and estimate. Other important moments defined in the treatment effects literature, like the Average Treatment Effect (ATE) or the Local Average Treatment Effect (LATE) in the context of instrumental variables, can be derived from the MTE curve by integrating the MTE curve using appropriate weights.

Furthermore, the shape of the MTE function with respect to \( v \) illustrates how the average treatment effects changes as the unobserved resistance to treatment varies. For example, if the MTE function is a decreasing function of \( v \), the pattern of selection into treatment is consistent with a Roy model in which there is positive selection into treatment, that is, individuals that are more likely to self-select into treatment have higher average treatment effects.

### 6.1 Identification of the MTE function with a continuous instrument \( Z \)

Again, the critical issue in the identification of the MTE function is that the potential outcomes \( Y_0 \) and \( Y_1 \) are not simultaneously observed for a given household. The econometrician only gets to observe a random sample of \( (Y, X, Z, D) \), where \( Y \) depends on the self-selected treatment choice.

Heckman and Vytlacil (1999) have shown sufficient conditions for the nonparametric identification of the MTE when the instrumental variable \( Z \) is continuous. In a very similar fashion to the identification argument of LATE (Imbens and Angrist, 1994), the key identification assumption of the MTE relies on an exclusion restriction of \( Z \): \( Z \) is included in the selection equation but is excluded from the potential outcome equations.

More formally, \( \text{MTE}(x, v) \) is identified if the following sufficient assumption holds:

**Assumption 1 Conditional independence of \( Z \): \( Z \perp \{U_0, U_1, V\} \)**

In words, this assumption implies that the instrumental variable \( Z \) only has an indirect effect on the potential outcomes \( Y_0 \) and \( Y_1 \) via its direct effect on the treatment choice \( D \), after controlling for the vector of observables \( X \). Assumption 1 also implies that \( Z \) is independent of the unobserved resistance to treatment \( V \), after controlling for \( X \). In Appendix A I present a brief identification argument summarized from (Heckman and Vytlacil, 2005) for the sake of completeness.

\( \text{MTE}(x, v) \) is identified locally if there are individuals with observed characteristics \( x \) and propensity score \( v \). However, the identification of other relevant population parameters like

These models also require extra assumptions on how childcare seats are allocated.
the average treatment effect of households with observed characteristics $x$, $ATE(x)$, which is defined as $ATE(x) = E[Y_1 - Y_0 \mid X = x]$, requires the integration of $MTE(x,v)$ over the full support of the unobserved resistance to treatment $V$, which has support $[0,1]$ by construction. Thus, the nonparametric identification of these other parameters require that the support of $P(X,Z)$ conditional on $X = x$ extends from 0 to 1.

In practice, this support requirement is infeasible in most cases. For example, the number of possible values that the vector of observable household characteristics $X$ can take increases exponentially as more regressors are included in the analysis. Thus, the empirical literature that has estimated MTE functions and other population parameters has adopted the following auxiliary assumption that relaxes the support requirements:

**Assumption 2: Additive separability of the MTE function:** $E[U_1 - U_0 \mid X,V] = E[U_1 - U_0 \mid V]$

If Assumption 2 holds, then the equation of $MTE(x,v)$ simplifies to the following expression:

$$MTE(x,v) = \mu_1(x) - \mu_0(x) + E[U_1 - U_0 \mid V = v] \quad (6)$$

Hence, under this assumption $MTE(x,v)$ becomes additively separable on $x$ and $v$. This implies that the shape of the MTE curve with respect to $v$ does not depend on $x$. A change in the observed characteristics $x$ will only shift up or down the MTE curve. Furthermore, under Assumptions 1 and 2, the identification of important population parameters like $ATE(x)$, $ATT(x)$, or $ATUT(x)$ only requires that the unconditional support of $P(X,Z)$ spans from 0 to 1.

Even if the additive separability of the MTE function might be a restrictive assumption in practice, many papers that have empirically implemented the MTE framework have assumed this assumption for tractability (Brinch et al., 2017; Carneiro et al., 2011; Cornelissen et al., 2018; Fefe and Laliv, 2018a). I also assume that the MTE function is additively separable.

## 7 Empirical implementation of MTE

### 7.1 Definition of treatment $D$ and covariates $X$

Based on the empirical findings that the effects of expanding childcare centers for children ages 0–3 at the local level are much stronger for mothers than for fathers, I focus on the household choice about the duration of maternity leave. Again, I define that a household $i$ is treated ($D_i = 1$), if the duration of maternal leave is at most 14 months.

From the set of observed household characteristics $X$, I separate the location of residence ($L$) and the birth cohort of the child ($C$), from the rest of the observed characteristics of the households ($\bar{X}$). Thus, I rewrite $X$ as $X = (\bar{X}, L, C)$. In my baseline specification of MTE $\bar{X}$ contains all the demographic controls included in my preferred specification of the generalized difference-in-differences explained in Section 4: the level of education of the mother, the age of the mother at
the time of birth, and the two foreign-born indicators for the mother. I now also include the age of the child at the time of the interview as an important regressor.  

### 7.2 Outcomes

I focus on two different sets of outcomes: outcomes related to the labor market attachment of mothers and outcomes related to maternal earnings and social benefits. For the first type of outcomes, I analyze the marginal treatment effect of short leaves on the probability that mothers are working and on the actual number of hours worked per week at the time of the interview. For the second set of outcomes, I study the marginal treatment effects of short leaves on the maternal monthly income and the probability that the household is receiving transfer payments.

My empirical implementation of MTE is closest to Schuss and Azaouagh (2021) and Yamaguchi et al. (2018), which are two recent papers that estimate the MTE of attendance in early childcare on parental labor market outcomes in Germany and Japan, respectively. Different to these papers, my treatment variable is defined in terms of leave duration. Mothers in Germany inform their employers about their leave plans during the first week after the child is born. In contrast, attendance in early childcare is usually only possible after the child turns two months old. Hence, I argue that decisions about parental leave precede childcare attendance outcomes. Furthermore, all my outcomes of interest are measured after the child turns three, when all children in the sample have a guaranteed seat in kindergarten. In contrast, Schuss and Azaouagh (2021) and Yamaguchi et al. (2018) focus on the concurrent effects of childcare attendance on labor market outcomes.

### 7.3 Choice of Z and discussion about the conditional independence assumption

The natural choice to use as an instrument $Z$ is the number of nurseries and Kitas in the locality per 1,000 children in the age group 0–3. I have previously shown in Table 3 that the normalized number of nurseries and Kitas in the locality has a substantially and statistically significant effect on the take-up of short leaves. Additionally, the comparison between columns (1) and (2) of Table 3 and the estimates from Table 6 provide suggestive evidence that this measure of childcare availability appears to be uncorrelated with observed household characteristics, after controlling for locality and birth cohort fixed effects. Given that $Z$ appears to be conditionally uncorrelated with observed characteristics $\tilde{X}$, it is then also plausible that $Z$ is uncorrelated with $V$, the unobserved resistance to treatment.

The assumption of conditional independence of $Z$ also requires that $Z \perp \{U_0, U_1\}$. In my particular set-up, the conditional independence assumption only allows the local availability of childcare to affect subsequent maternal labor market outcomes through its effect on the duration of

---

33The inclusion of the age of the focal child at the time of the interview will help distinguish whether the effects of short leaves relative to long leaves fade out or grow in importance over time.

34I define that mothers are working if they report to be working at least 10 hours per week. Less than 5% of mothers that report to be working are working less than 10 hours per week.

35As briefly mentioned in Section 3, I define that a household is receiving transfer payments at the time of the interview if the respondent reports that any member of the household is receiving any of the following social benefits: short-term or long-term unemployment, social assistance (Sozialhilfe), or housing benefits (Wohngeld).
maternal leave. This exclusion restriction is consistent with two key causal mechanisms highlighted in the literature on female labor participation: human capital depreciation and habit persistence.

However, the availability of childcare in the year of birth of the child might potentially affect maternal labor market outcomes when the child is older through different mechanisms that do not exclusively rely on the time out of work. I highlight four of these alternative channels: human capital accumulation of the child, intra-household bargaining power, outside options in the labor market, and general equilibrium effects.

Early childcare attendance in Germany has been shown to have heterogenous effects on the cognitive and noncognitive skills of children (Felfe and Lalive, 2018a). A first violation of the exclusion restriction would happen if households consider the human capital of the child when deciding about the employment status of the mother when the child is old enough to attend kindergarten. I argue that this concern is not first-order in the case of Germany. Kindergarten attendance by 2007 was already above 90% in the majority of localities. In this sense, most households were already sending their children to kindergarten at the start of the analysis period, regardless of their demographic background, actual attendance to early childcare, and prior maternal labor force decisions.36

In terms of the intra-household bargaining power mechanism, I did find a significant negative effect of increasing childcare availability in the locality on the probability that the focal child is observed in an intact household in subsection 5.4. However, this effect was economically modest. Additionally, I found in the same subsection 5.4 that the effect of childcare availability on paternity leave outcomes is a precisely estimated zero. These results provide suggestive evidence that the effects of childcare availability on intra-household bargaining power is limited.

I am not able to rule out the last two mechanisms entirely. An increase in the availability of childcare might enhance the outside options of working mothers, for example, by modifying the set of workplaces that mothers perceive as desirable. In terms of general equilibrium effects, firms might change their promotion and compensation policies given the increased availability of childcare in the locality. However, the inclusion of locality and birth cohort fixed effects in the vector of observed characteristics X partially captures the effects of these two mechanisms.

### 7.4 Parametrization of the potential outcomes equations

The exogenous source of variation in the data stems from the differences in the roll-out of the childcare expansion across locations. Hence, I also include locality and birth cohort fixed effects in the potential outcomes equations. I further assume that the potential outcome equations (4) is linear with respect to the observed household characteristics \( x_i = (\bar{x}_i, l_i, c_i) \) as follows:

---

36The proportion of Kindergärten and Kitas that offer full-time services has grown over time and varies substantially across localities (Felle and Lalive, 2018b). In robustness checks, I add as an additional time-varying regressor the proportion of children ages 3-6 that attend a childcare in a full-time manner in the locality.
\[
Y_{d,i}(\tilde{x}_i, l_i, c_i, U_{d,i}) = \beta_{x,d}\tilde{x}_i + \sum_{\ell \in L} \beta_{\ell,d} 1 (l_i = \ell) + \sum_{c \in C} \beta_{c,d} 1 (c_i = c) + U_{d,i} \quad d \in \{0, 1\}
\] (7)

However, in order to reduce the number of parameters to estimate as part of the MTE curve I follow the empirical implementation of MTE of Cornelissen et al. (2018) and Felfe and Lalive (2018a) that assumes that the effects of a given location or birth cohort on outcomes do not vary by treatment status, which means that \( \beta_{\ell,0} = \beta_{\ell,1} \forall \ell \in L \) and \( \beta_{c,0} = \beta_{c,1} \forall c \in C \). Based on these simplifying assumptions, the location and cohort fixed effects cancel out in the MTE equation and equation (6) can be rewritten as:

\[
MTE(\tilde{x}_i, l_i, c_i, v_i) = (\beta_{x,1} - \beta_{x,0}) \tilde{x}_i + E[ U_1 - U_0 | V = v_i ]
\] (8)

Even though the identification result in Appendix A shows that \( MTE(\tilde{x}_i, l_i, c_i, v_i) \) is nonparametrically identified, I assume that \( k(v_i) \) is a second-order polynomial in my baseline specification. However, I perform robustness checks in which I vary the order of the polynomial \( k(v_i) \). I also include estimates of the MTE function using a nonparametric approach.

### 7.5 Parametrization of the treatment selection equation

I parametrize the treatment selection equation shown in (5) by assuming that the function \( \mu_D \) is linear with respect to the parameters of \( X \). Following Carneiro et al. (2011) among others, I also include interaction terms between the instrument \( Z \) and each one of the household covariates in \( \tilde{X} \) to ease the estimation of the vector \( \beta_{x} \) in equation (8). Then, treatment selection equation (5) can be rewritten as:

\[
D_i(\tilde{x}_i, l_i, c_i, z_i, U_{D,i}) = 1 \left( \gamma_x \tilde{x}_i + \gamma_z z_i + \gamma_x \tilde{x}_i z_i + \sum_{\ell \in L} \gamma_\ell 1 (l_i = \ell) + \sum_{c \in C} \gamma_c 1 (c_i = c) + U_{D,i} \geq 0 \right) \quad d \in \{0, 1\}
\] (9)

Finally, I assume that \( U_D \) is distributed as a standard normal distribution. Hence, I estimate the vector of first stage parameters \( \gamma \) via probit.

### 7.6 Estimation method

I estimate the parameters of the MTE using a two-stage procedure. In the first stage, I estimate the parameters of the treatment selection equation and the propensity score for each household via probit. Based on the support of the propensity score of treated and untreated individuals, I redefine the common support of the propensity score in a conservative fashion by additionally dropping 1% of the treated and untreated sample in the tails of the original common support.
following Brinch et al. (2017).

In the second stage, the MTE parameters are estimated using the local IV approach, which obtains the numerical derivative of \( E[Y \mid X, P(D = 1 \mid X, Z)] \) with respect to the propensity score in the region of common support (see Appendix A for an explanation). The standard errors of the MTE estimates are obtained using 100 bootstrapped iterations, in order to consider that the propensity score used in the second stage was estimated from the probit model in the first stage.\(^{37}\)

## 8 Results from the MTE framework

### 8.1 Treatment selection equation

I have already presented results of a probit specification that links the probability to take a short leave (which corresponds to treatment branch \( D = 1 \) in the MTE framework) to the local availability of childcare and observed household characteristics in Table 3. Based on column (2) of Table 3, an additional nursery or Kita per 1,000 children increases the average take-up of short leaves by 1.6 percentage points.

The only difference between the treatment selection equation and the probit estimated in Table 3 is that the treatment selection equation includes interaction terms between the normalized number of childcare centers and the rest of the demographic characteristics of the household. In Appendix Table A-2 I show the estimated average marginal effects of the local availability of childcare for children ages 0–3 on the take-up of short leaves using the probit specification in (9).

The baseline specification is shown in column (1). It includes the same set of household covariates as the preferred specification in the generalized difference-in-differences section and the age of the kid at the time of the interview. The estimated average marginal effect of an additional childcare center per 1,000 children in the age group 0–3 is 1.8 percentage points, which is very close to the estimate of 1.6 percentage points that was obtained from the probit specification without interaction terms.

In columns (2)–(5) of Appendix Table A-2 I include additional regressors as robustness checks. Column (2) includes the local unemployment rate and the number of births per women in the age group 15–45 in the year of birth of the child. Column (3) additionally includes the childcare quality index in the year of birth. Column (4) adds two time-varying characteristics at the local level in the year when the child turns 3 years old: the proportion of children in the age range 3 – 6 that attends kindergarten in a full-time manner and an indicator variable on whether kindergarten is free when the child is 3 years old. Finally, column (5) contains the same demographic characteristics as column (1) but also includes the number of children in the household at the time of the interview. Across columns (2)–(5), the estimated average marginal effect of childcare availability on the take-up of short leaves is quite similar to the estimated average marginal effect shown in column (1). Importantly, the p-values associated to the hypothesis test that the marginal effect of childcare availability is equal to zero are all smaller than 0.01.

\(^{37}\)In practice, I base most of my code on the pre-programmed command \textit{mitefe} in Stata (Andresen, 2019).
8.2 Propensity scores and common support

The propensity score for each household is calculated using the estimates of column (1) of Appendix Table 3. In Figure 6 I show how the estimated propensity scores vary as the instrument Z increases, by the type of school-leaving certificate of the mother. An important take-away from Figure 6 is that the gap in the take-up of short leaves by the education of the mother narrows down as the local availability of childcare increases. Figure 6 is also consistent with the finding in subsection 5.5 that the leave decisions of mothers with low levels of education are more sensitive to changes in childcare availability. This might be a result of the childcare allocation system that prioritizes mothers from advantaged backgrounds when the rationing of seats is more extreme. However, this difference might also originate from an insurance motive of families that come from a more disadvantaged background. If the risk to not secure a childcare spot is sufficiently high, parents without access to informal caregivers or without monetary resources to secure a private nanny or au-pair might opt to take a long leave as insurance.

The original common support of the propensity score between the untreated and treated individuals extends from 0.29 to 0.94. After trimming 1% of the untreated and treated sample in the tails of the original common support distribution, the conservative common support that I will use to compute rescaled different population treatment effects spans from 0.36 to 0.87. Although far from unity, the size of the common support is not too far from other papers in the MTE literature.

The empirical distributions of the propensity score for the treated and non-treated individuals are shown in Figure 7. Panel A presents the empirical distributions of the propensity score when only the normalized number of Kitas and nurseries in the locality (Z) is included as an instrument. Instead, Panel B shows the empirical distributions of the propensity score when the interaction terms between the local availability of childcare and the observed household characteristics are included as instruments. Figure 7 shows that the common support of the propensity score is virtually unchanged when the interaction terms are included as instruments.

8.3 Returns to treatment based on observed characteristics

Table 7 presents the estimated coefficients of the vector $\beta_{x,0}$, which corresponds to the vector of parameters of the potential outcomes equation in (7) when $D = 0$, and $\beta_x$, which appears in the expression of the simplified MTE curve in (8). In each column of Table 7, I present the estimates of $\beta_{x,0}$ and $\beta_x$ for a different labor market outcome. Column (1) shows the estimates for the indicator variable that the mother is working at least 10 hours per week; column (2), the number of hours that the mother works per week; column (3), the monthly income of the mother; and column (4), an indicator variable on whether the household is receiving transfer payments.

A clear picture with respect to the age of the mother at the time of birth emerges across all outcomes in Table 7. In the untreated state, which corresponds to the take-up of a long leave, the age of the mother at birth does not seem to play a statistically significant role on outcomes.

---

38A similar result is found by Jessen et al. (2020) in terms of childcare attendance. The authors find that the education gap in childcare attendance in a given locality narrows down as the excess demand for childcare decreases in the locality.
However, the estimated coefficient for this same variable in the vector $\beta_x$ is statistically significant and positive for the work and income outcomes and negative for the indicator variable on transfer receipt. This means that the average effect of taking a short leave instead of a long leave on the probability to work, the number of hours worked per week, and the monthly income of the mother are statistically significantly higher for older women. In contrast, in the group of mothers that took short leaves older mothers are significantly less likely to claim transfer payments than younger mothers.

The total experience of the mother prior to birth is unfortunately unobserved in the KiBS dataset. However, the age of the mother at the time of birth acts as a proxy for this variable. The results that the relative benefits of short leaves are higher for mothers with more accumulated experience is in line with the atrophy estimates in Adda et al. (2017). These authors find that the annual depreciation rates of human capital are higher after 6 years of uninterrupted work experience than after 3 years of uninterrupted work experience. Finally, the result that the estimates of age at the time of birth in the vector $\beta_{x,0}$ are small and not statistically significant is consistent with a story in which long leaves erase most of the benefits brought by the accumulated experience prior to birth.

In terms of the nationality of the mother, the estimates are noisier. According to the estimates from Table 7, foreign-born mothers are more likely to be receiving transfer payments and less likely to be working at the time of the interview after a long leave. However, the take-up of short leaves helps to narrow these outcome gaps by nationality status. In the case of transfer receipt, the gap across national groups is entirely erased. For example, the gap in transfer payment receipt between mothers born in Eastern Europe and native mothers is equal to 17% in the untreated state (after a long leave). However, the gap in the treated state (after a short leave) is equal to $17\% - 21\% = -4\%$.

Finally, the coefficients on the age of the kid at the time of the interview help identify whether the effects of the treatment on subsequent labor market outcomes fade out, persist, or grow in importance over time. After the take-up of a long leave, the number of years that have passed since the focal child was born has a positive effect on the work and income variables. The negative coefficient of this same variable in the vector $\beta_x$ implies that the slope of the average labor market outcomes as a function of time after birth is smaller after a short leave than after a long leave. This provides suggestive evidence that the positive effects of short leaves on subsequent labor market outcomes fade out over time.

The p-value associated to the null hypothesis that the effects of a short leave do not vary by observed household characteristics is included in Table 7. This test corresponds to testing that the vector $\beta_x$ is equal to zero. For all four outcomes, the null hypothesis can be rejected at conventional levels.

## 8.4 Returns to treatment based on unobserved characteristics

With the estimates of $\beta_x$ and $\gamma$, I can find the MTE curve described in equation (8) for each one of the four different outcomes. Figure 8 shows the estimated MTE curves as a function of the
unobserved resistance to treatment $V$ for each of the four outcomes.\footnote{Each of the curves has been estimated at the sample mean of the observed household characteristics $\bar{X}$. Given that the MTE function is additively separable, the shape of the curve is independent of $X$.}

Panels A to C of Figure 8 provide evidence that the relationship between the unobserved resistance to treatment $V$ and the average unobserved gains from treatment has a U-shape. This implies that the mothers that are the most likely or least likely to take a short leave based on their unobserved characteristics are the mothers that have the highest effects of the treatment in terms of their subsequent labor market outcomes.

Table 7 also shows the p-values associated with the hypothesis test that the vector of parameters of the polynomial $k(v)$ are equal to zero ($\gamma = 0$). I can reject the null hypothesis of no essential heterogeneity for the indicator variable that the mother is working at least 10 hours per week and the number of hours worked per week at the time of the interview at conventional significance levels. Unfortunately, the estimates for monthly maternal income are somewhat noisier. The p-value associated to the test of no essential heterogeneity is equal to 0.157 for monthly income. I also fail to reject the null that the unobserved returns to treatment on transfer receipt are homogenous.

The nonlinear relationship between the unobserved resistance to treatment and the average returns to treatment is indicative that the selection pattern of mothers into short leaves compared to long leaves is highly complex. In particular, a simple Roy model that only allows for positive or negative patterns of selection into short leaves would not be able to replicate the empirical patterns of selection shown in Panels A through C. Furthermore, even fully dynamic discrete choice models that include different wage or taste fixed effects might have trouble reconciling the U-shaped patterns of selection, especially in terms of maternal income. This is the case because most types of fixed effects would not interact with the duration of leave, and would thus cancel out from the computation of the MTE equation.

### 8.5 Estimates of ATE, ATT, ATUT, and LATE

In Table 8 I show the estimates of re-scaled versions of the average treatment effect (ATE), the average treatment effect on the treated (ATT), and the average treatment effect on the untreated (ATUT) of the take-up of short leaves relative to long leaves on the four different labor market outcomes.\footnote{The ATE, ATUT, and ATT can be calculated by integrating the estimated MTE curves with the weights derived in \textit{Heckman and Vytlacil} (1999). As the common support in my empirical specification is not the interval $[0, 1]$, I rescale the weights in the common support so that they integrate to 1. This is a common practice in the empirical implementation of MTE.} Importantly, the LATE associated with the instrumental variable $Z$, the normalized number of nurseries and Kitas in the locality in the year of birth of the focal child, can also be calculated from the MTE curve by using the weights derived in \textit{Heckman and Vytlacil} (1999). I include the estimates of the LATE associated to the German childcare expansion in Table 8.

Table 8 shows that the point estimates of the ATE of short leaves compared to long leaves is positive for the work and income outcomes and negative for the indicator of transfer receipt. Although the standard errors associated of the ATE of some of the outcomes are somewhat large, the point estimates provide suggestive evidence that short leaves do have positive average causal
effects on subsequent labor market outcomes. In the case of the number of hours worked per week at the time of the interview, the ATE is equal to 14.69 hours and is statistically significant at the 5% level. This result indicates that the take-up of short leaves relative to long leaves indeed promotes an average increase in the number of weekly hours that mothers work once that the child is older.

Importantly, the shape of the MTE curve also plays a crucial role in assessing whether the Local Average Treatment Effect (LATE) identified by a particular policy change can be extrapolated beyond the sub-population of compliers. Table 8 shows that the LATE from the German childcare expansion underestimates the ATE, ATT, and ATUT of the take-up of short leaves on subsequent labor market outcomes. This is the case because the weights assigned to calculate the LATE give more importance to the individuals in the middle region of the unobserved resistance to treatment. This set of individuals correspond to the individuals with the lowest unobserved returns to treatment.

9 Conclusion

In this paper I have analyzed whether household decisions about parental leave react to plausibly exogenous changes in the local availability of subsidized childcare. I have found robust evidence that the take-up of long maternal leaves decreases as the local availability of subsidized childcare increases, which indicates that households see the entitlement to a long maternal leave and subsidized childcare as substitutes. By contrast, I do not find any evidence that the behavior of fathers is affected by changes in the provision of subsidized childcare.

Ideally, a cost-benefit analysis of a childcare expansion should carefully consider the counterfactual employment status of mothers when children are infants or toddlers in the scenario with and without the childcare expansion.41 My results show that the childcare expansion shifted a considerable proportion of mothers that would have taken a long leave in the counterfactual scenario without childcare expansion to return to work earlier in the scenario with childcare expansion.

The potential reduction in the firms’ costs from shorter leaves has to be considered in an accurate cost-benefit analysis of an expansion in childcare.42 These savings are potentially higher if individuals are entitled to long job-protected leaves, which would suggest that current cost-benefit analyses systematically underestimate the benefits of expanding subsidized childcare. Furthermore, the underestimation in the benefits of expanding childcare is potentially higher in countries with long job-protected leaves. I leave the task of carefully assessing all the costs and benefits of the German childcare expansion for future work.

In this paper I also adopt a MTE framework to show that the take-up of a short maternal leave relative to a long maternal leave has overall positive effects on subsequent maternal labor market outcomes, including hours worked and monthly income. Furthermore, I conclude that the effects of this treatment are heterogenous in terms of observed characteristics (especially in terms of the

41This claim is analogous to the argument in Kline and Walters (2016) that an accurate cost-benefit analysis of Head Start needs to consider the reduction in costs associated to a lower attendance in other publicly subsidized programs.
42Ginja et al. (2020) provides one of the few available estimates of the costs associated with longer leaves for firms.
age of the mother at the time of birth) and unobserved characteristics. The empirical finding that the shape of the average returns to treatment as a function of the unobserved resistance to treatment has a U-shape suggests that policies that target mothers that are unlikely to take short leaves can potentially have larger effects (e.g., Kuka and Shenhav (2020)).

Although the MTE framework is quite flexible, it does restrict the way in which the local availability of childcare at the time of birth can affect subsequent labor market outcomes. Another potential mechanism through which the local availability of childcare affects the labor market outcomes of mothers is by directly modifying their outside options (Caldwell and Danieli, 2020). This might be the case if the decision of where to work is contingent on the presence of a childcare center near the workplace. Insufficient childcare provision might also help explain why women have a lower willingness to commute longer distances to work than men (Le Barbanchon et al., 2021). A promising venue for future work is to analyze how expansions in the local availability of childcare affect outside options and job search behavior of parents at a finer geographical level.
References


M. Andresen. Mtefe: Stata module to compute marginal treatment effects with factor variables. 2019. 31


M. Thomas. The impact of mandated maternity leave policies on the gender gap in promotions: Examining the role of employer-based discrimination. *Available at SSRN 3729663*, 2020. 4


### Table 1: Summary statistics

<table>
<thead>
<tr>
<th></th>
<th>All</th>
<th>Short leave</th>
<th>Long leave</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal leave duration</td>
<td>17.141</td>
<td>11.205</td>
<td>27.548</td>
</tr>
<tr>
<td></td>
<td>(9.215)</td>
<td>(2.222)</td>
<td>(7.419)</td>
</tr>
<tr>
<td>1(Short leave)</td>
<td>0.637</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>(0.481)</td>
<td>(0)</td>
<td>(0)</td>
</tr>
<tr>
<td>1(Mother passed Abitur)</td>
<td>0.587</td>
<td>0.642</td>
<td>0.491</td>
</tr>
<tr>
<td></td>
<td>(0.492)</td>
<td>(0.479)</td>
<td>(0.500)</td>
</tr>
<tr>
<td>1(Foreign mother, Eastern Europe)</td>
<td>0.056</td>
<td>0.049</td>
<td>0.070</td>
</tr>
<tr>
<td></td>
<td>(0.231)</td>
<td>(0.215)</td>
<td>(0.255)</td>
</tr>
<tr>
<td>1(Foreign mother, not Eastern Europe)</td>
<td>0.051</td>
<td>0.051</td>
<td>0.051</td>
</tr>
<tr>
<td></td>
<td>(0.219)</td>
<td>(0.219)</td>
<td>(0.219)</td>
</tr>
<tr>
<td>Age of mother at birth</td>
<td>32.735</td>
<td>32.563</td>
<td>33.038</td>
</tr>
<tr>
<td></td>
<td>(4.455)</td>
<td>(4.443)</td>
<td>(4.460)</td>
</tr>
<tr>
<td>Birth cohort year</td>
<td>2011.5</td>
<td>2011.6</td>
<td>2011.3</td>
</tr>
<tr>
<td></td>
<td>(1.765)</td>
<td>(1.761)</td>
<td>(1.757)</td>
</tr>
<tr>
<td>Local # of nurseries and Kitas (per 1,000 children in ages 0–3)</td>
<td>15.239</td>
<td>15.906</td>
<td>14.071</td>
</tr>
<tr>
<td></td>
<td>(5.333)</td>
<td>(5.467)</td>
<td>(4.876)</td>
</tr>
<tr>
<td>At the time of household interview:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1(Biological father present)</td>
<td>0.920</td>
<td>0.919</td>
<td>0.921</td>
</tr>
<tr>
<td></td>
<td>(0.272)</td>
<td>(0.273)</td>
<td>(0.270)</td>
</tr>
<tr>
<td># of children in household</td>
<td>1.767</td>
<td>1.745</td>
<td>1.805</td>
</tr>
<tr>
<td></td>
<td>(0.423)</td>
<td>(0.436)</td>
<td>(0.396)</td>
</tr>
<tr>
<td>Age of the focal child (months)</td>
<td>60.26</td>
<td>59.41</td>
<td>61.77</td>
</tr>
<tr>
<td></td>
<td>(18.96)</td>
<td>(18.85)</td>
<td>(19.04)</td>
</tr>
<tr>
<td>1(Mother working ≥ 10 hours)</td>
<td>0.746</td>
<td>0.797</td>
<td>0.656</td>
</tr>
<tr>
<td></td>
<td>(0.436)</td>
<td>(0.402)</td>
<td>(0.475)</td>
</tr>
<tr>
<td>Maternal hours worked</td>
<td>22.46</td>
<td>25.48</td>
<td>17.14</td>
</tr>
<tr>
<td></td>
<td>(14.42)</td>
<td>(14.20)</td>
<td>(13.24)</td>
</tr>
<tr>
<td>1(Household receiving transfers)</td>
<td>0.064</td>
<td>0.050</td>
<td>0.089</td>
</tr>
<tr>
<td></td>
<td>(0.244)</td>
<td>(0.217)</td>
<td>(0.285)</td>
</tr>
<tr>
<td>Monthly maternal income</td>
<td>1,216.6</td>
<td>1,415.9</td>
<td>863.0</td>
</tr>
<tr>
<td></td>
<td>(998.6)</td>
<td>(1,045.3)</td>
<td>(795.5)</td>
</tr>
<tr>
<td>Number of observations</td>
<td>12,643</td>
<td>8,051</td>
<td>4,592</td>
</tr>
</tbody>
</table>

Note: Table 1 presents summary statistics of the main sample. For a description of the main sample, see Section 3. The first column shows the summary statistics for all the households in the sample. The second (third) column shows the summary statistics for the sub-sample of households where maternal leave was 14 months or shorter (15 months or longer).
Table 2: Effects of local childcare availability on maternal leave duration, all Germany

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of nurseries and Kitas (per 1,000 children ages 0–3)</td>
<td>-0.411</td>
<td>-0.429</td>
<td>-0.434</td>
<td>-0.453</td>
</tr>
<tr>
<td></td>
<td>(0.076)</td>
<td>(0.077)</td>
<td>(0.106)</td>
<td>(0.102)</td>
</tr>
<tr>
<td>1(Mother passed Abitur)</td>
<td>-2.760</td>
<td>-2.815</td>
<td>-2.815</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.301)</td>
<td>(0.290)</td>
<td>(0.290)</td>
<td></td>
</tr>
<tr>
<td>1(Foreign mother, Eastern Europe)</td>
<td>0.781</td>
<td>0.780</td>
<td>0.780</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.430)</td>
<td>(0.430)</td>
<td>(0.430)</td>
<td></td>
</tr>
<tr>
<td>1(Foreign mother, not Eastern Europe)</td>
<td>-0.515</td>
<td>-0.517</td>
<td>-0.517</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.337)</td>
<td>(0.337)</td>
<td>(0.337)</td>
<td></td>
</tr>
<tr>
<td>Age of mother at birth</td>
<td>0.126</td>
<td>0.128</td>
<td>0.128</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td>(0.017)</td>
<td>(0.017)</td>
<td></td>
</tr>
<tr>
<td>Local unemployment rate</td>
<td>0.039</td>
<td>0.026</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.230)</td>
<td>(0.221)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local number of births per women ages 15–45</td>
<td>0.019</td>
<td>0.025</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.061)</td>
<td>(0.061)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Index of childcare quality</td>
<td>3.183</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(2.050)</td>
<td></td>
</tr>
<tr>
<td>Locality FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Birth cohort FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Mean maternal leave duration (in months)</td>
<td>17.14</td>
<td>17.14</td>
<td>17.14</td>
<td>17.14</td>
</tr>
<tr>
<td>F-statistic of $H_0 : \beta = 0$</td>
<td>29.21</td>
<td>31.22</td>
<td>16.81</td>
<td>19.52</td>
</tr>
<tr>
<td>p-value of $H_0 : \beta = 0$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Number of observations</td>
<td>12,644</td>
<td>12,644</td>
<td>12,644</td>
<td>12,644</td>
</tr>
</tbody>
</table>

Note: Table 2 presents the estimated coefficients of the generalized difference-in-differences shown in Equation 1. All regressions are estimated by Ordinary Least Squares. The dependent variable is the maternal leave duration in months. The main independent variable is the number of nurseries and Kits per 1,000 children in age group 0–3 in the locality. All specifications include locality and birth cohort fixed effects. The specification in column (2) includes observed household characteristics. The specification in column (3) additionally controls for the local unemployment rate and number of births per women in the age group 15–45. Finally, column (4) also includes an index of childcare quality at the local level. All standard errors are clustered at the locality level.
### Table 3: Marginal effects of local childcare availability on take-up of short leaves, all Germany

<table>
<thead>
<tr>
<th>Probit</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>LPM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of nurseries and Kitas</td>
<td>0.015</td>
<td>0.016</td>
<td>0.022</td>
<td>0.023</td>
<td>0.019</td>
</tr>
<tr>
<td>(per 1,000 children ages 0–3)</td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>(1) (Mother passed Abitur)</td>
<td>0.131</td>
<td>0.131</td>
<td>0.131</td>
<td>0.134</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.011)</td>
<td>(0.011)</td>
<td>(0.013)</td>
<td></td>
</tr>
<tr>
<td>(1) (Foreign mother, Eastern Europe)</td>
<td>0.024</td>
<td>0.024</td>
<td>0.024</td>
<td>0.024</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.018)</td>
<td>(0.018)</td>
<td>(0.018)</td>
<td>(0.018)</td>
<td></td>
</tr>
<tr>
<td>(1) (Foreign mother, not Eastern Europe)</td>
<td>-0.038</td>
<td>-0.038</td>
<td>-0.038</td>
<td>-0.039</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.019)</td>
<td>(0.019)</td>
<td>(0.020)</td>
<td>(0.021)</td>
<td></td>
</tr>
<tr>
<td>Age of mother at birth</td>
<td>-0.006</td>
<td>-0.006</td>
<td>-0.006</td>
<td>-0.006</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td></td>
</tr>
<tr>
<td>Local unemployment rate</td>
<td>-0.023</td>
<td>-0.023</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td>(0.009)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local number of births per</td>
<td>-0.002</td>
<td>-0.002</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>women ages 15–45</td>
<td>(0.004)</td>
<td>(0.009)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Index of childcare quality</td>
<td>-0.205</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.101)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Locality FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Birth cohort FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Mean dependent variable</td>
<td>0.637</td>
<td>0.637</td>
<td>0.637</td>
<td>0.637</td>
<td>0.637</td>
</tr>
<tr>
<td>(\chi^2/F)-statistic of (H_0: \beta = 0)</td>
<td>14.19</td>
<td>16.28</td>
<td>16.29</td>
<td>17.07</td>
<td>19.99</td>
</tr>
<tr>
<td>p-value of (H_0: \beta = 0)</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Number of observations</td>
<td>12,642</td>
<td>12,642</td>
<td>12,642</td>
<td>12,642</td>
<td>12,644</td>
</tr>
</tbody>
</table>

Note: Table 3 presents the estimated average marginal effect of a change in the number of nurseries and Kitas per 1,000 children in the age group 0–3 on the probability that mothers take “short” leaves. Columns (1) – (4) present estimates from probit specifications; column (5) is a linear probability model. All specifications include locality and birth cohort fixed effects. The specifications in columns (2) and (5) include observable household characteristics. The specification in column (3) additionally controls for the local unemployment rate and number of births per women ages 15–45 in the year of birth. Finally, column (4) also includes an index of childcare quality in the locality. All standard errors are clustered at the locality level.
Table 4: Effect of local childcare availability on other outcomes

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Intact household)</td>
<td>Maternal leave</td>
<td>(Paternal leave &gt; 0)</td>
<td>Paternal leave</td>
</tr>
<tr>
<td>Number of nurseries and Kitas (per 1,000 children ages 0–3)</td>
<td>-0.006</td>
<td>-0.484</td>
<td>0.005</td>
<td>0.043</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.083)</td>
<td>(0.007)</td>
<td>(0.027)</td>
</tr>
<tr>
<td>I(Mother passed Abitur)</td>
<td>0.045</td>
<td>-2.716</td>
<td>0.152</td>
<td>0.551</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.294)</td>
<td>(0.010)</td>
<td>(0.062)</td>
</tr>
<tr>
<td>I(Foreign mother, Eastern Europe)</td>
<td>0.015</td>
<td>0.104</td>
<td>-0.061</td>
<td>-0.117</td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td>(0.527)</td>
<td>(0.018)</td>
<td>(0.145)</td>
</tr>
<tr>
<td>I(Foreign mother, not Eastern Europe)</td>
<td>-0.018</td>
<td>-0.502</td>
<td>-0.034</td>
<td>-0.041</td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td>(0.358)</td>
<td>(0.025)</td>
<td>(0.149)</td>
</tr>
<tr>
<td>Age of mother at birth</td>
<td>0.003</td>
<td>0.126</td>
<td>-0.002</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.018)</td>
<td>(0.001)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>Age of focal child at the time of the survey</td>
<td>-0.002</td>
<td>-0.002</td>
<td>-0.001</td>
<td>-0.005</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.018)</td>
<td>(0.001)</td>
<td>(0.006)</td>
</tr>
<tr>
<td>I(Father has abitur)</td>
<td>0.087</td>
<td>0.060</td>
<td>0.247</td>
<td>-0.002</td>
</tr>
<tr>
<td></td>
<td>(0.207)</td>
<td>(0.012)</td>
<td>(0.073)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>I(Foreign father, Eastern Europe)</td>
<td>-0.408</td>
<td>-0.090</td>
<td>0.035</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>(0.406)</td>
<td>(0.024)</td>
<td>(0.184)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>I(Foreign father, not Eastern Europe)</td>
<td>1.408</td>
<td>-0.099</td>
<td>-0.062</td>
<td>-0.002</td>
</tr>
<tr>
<td></td>
<td>(0.543)</td>
<td>(0.033)</td>
<td>(0.211)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>Locality FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Birth cohort FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Mean DV</td>
<td>0.921</td>
<td>17.080</td>
<td>0.487</td>
<td>1.627</td>
</tr>
<tr>
<td>p-value $H_0: \beta_Z = 0$</td>
<td>0.039</td>
<td>0.000</td>
<td>0.443</td>
<td>0.113</td>
</tr>
<tr>
<td>Number of observations</td>
<td>12,644</td>
<td>11,423</td>
<td>11,423</td>
<td>11,423</td>
</tr>
</tbody>
</table>

Note: Table 4 presents the estimated effects of a change in the number of nurseries and Kitas per 1,000 children in the age group 0–3 on alternative outcomes based on equation (1). Column (1) presents the effects of expanding childcare on the probability that the household is intact at the time of the survey for all households in the sample. Instead, the specifications in columns (2) to (4) limit the analysis to the sub-sample of intact households. The dependent variables in columns (2), (3), (4) correspond to the total duration of maternal leave, an indicator on whether the biological father took any paternal leave, and the total duration of paternal leave, respectively. All specifications are estimated via OLS and control for demographic characteristics and locality and birth cohort fixed effects. All standard errors are clustered at the locality level.
Table 5: Heterogeneous treatment effects of childcare expansion

Panel A: Level of education of the mother

<table>
<thead>
<tr>
<th></th>
<th>(1) All</th>
<th>(2) No Abitur</th>
<th>(3) Abitur</th>
<th>(2) – (3) Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of nurseries and Kitas (per 1,000 children ages 0–3)</td>
<td>-0.429 (0.078) [0.000]</td>
<td>-0.630 (0.129) [0.000]</td>
<td>-0.227 (0.099) [0.024]</td>
<td>-0.403 (0.163) [0.015]</td>
</tr>
<tr>
<td>Household characteristics</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Locality FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Birth cohort FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Mean D.V.</td>
<td>17.14</td>
<td>18.97</td>
<td>15.85</td>
<td></td>
</tr>
<tr>
<td>Number of observations</td>
<td>12,644</td>
<td>5,217</td>
<td>7,426</td>
<td></td>
</tr>
</tbody>
</table>

Panel B: Parity of the focal child

<table>
<thead>
<tr>
<th></th>
<th>(1) All</th>
<th>(2) Parity &gt; 1</th>
<th>(3) Parity = 1</th>
<th>(2) – (3) Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of nurseries and Kitas (per 1,000 children ages 0–3)</td>
<td>-0.460 (0.139) [0.002]</td>
<td>-0.603 (0.191) [0.003]</td>
<td>-0.192 (0.186) [0.324]</td>
<td>-0.411 (0.254) [0.116]</td>
</tr>
<tr>
<td>Household characteristics</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Locality FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Birth cohort FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Mean D.V.</td>
<td>17.06</td>
<td>17.94</td>
<td>15.85</td>
<td></td>
</tr>
<tr>
<td>Number of observations</td>
<td>9,188</td>
<td>5,331</td>
<td>3,857</td>
<td></td>
</tr>
</tbody>
</table>

Note: Table 5 presents heterogeneous treatment effects of local availability of childcare on maternal leave duration. Panel A presents heterogeneous treatment effects by the level of education of the mother, while Panel B shows heterogeneous treatment effects by the parity of the focal child. Column (1) in both panels shows the estimated effects for the full sample. Columns (2) and (3) of Panel A (B) show the estimated effects of childcare availability for subsets of the sample based on the level of maternal education (parity of the child). Column (4) presents the difference between the estimates in column (2) and column (3). All specifications are estimated via OLS and control for demographic characteristics and locality and birth cohort fixed effects. All standard errors are clustered at the locality level and shown in parentheses. p-value of the hypothesis test that the estimated coefficient is equal to zero is shown in brackets.
Table 6: Assessment of composition changes in different sub-samples

<table>
<thead>
<tr>
<th>D.V. Number of nurseries and Kitas (per 1,000 children ages 0–3)</th>
<th>(1) All</th>
<th>(2) Eligible</th>
<th>(3) Treated</th>
</tr>
</thead>
<tbody>
<tr>
<td>1(Mother passed Abitur)</td>
<td>0.012</td>
<td>-0.021</td>
<td>-0.053</td>
</tr>
<tr>
<td></td>
<td>(0.020)</td>
<td>(0.020)</td>
<td>(0.025)</td>
</tr>
<tr>
<td>1(Foreign mother, Eastern Europe)</td>
<td>0.009</td>
<td>0.013</td>
<td>-0.040</td>
</tr>
<tr>
<td></td>
<td>(0.029)</td>
<td>(0.034)</td>
<td>(0.039)</td>
</tr>
<tr>
<td>1(Foreign mother, not Eastern Europe)</td>
<td>-0.009</td>
<td>-0.018</td>
<td>-0.011</td>
</tr>
<tr>
<td></td>
<td>(0.029)</td>
<td>(0.034)</td>
<td>(0.039)</td>
</tr>
<tr>
<td>Age of mother at birth</td>
<td>-0.003</td>
<td>-0.000</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>Locality FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Birth cohort FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>p-value $H_0 : \beta_X = 0$</td>
<td>0.600</td>
<td>0.754</td>
<td>0.212</td>
</tr>
<tr>
<td>Number of observations</td>
<td>14,077</td>
<td>12,643</td>
<td>8,051</td>
</tr>
</tbody>
</table>

Note: Table 6 presents the estimated effects of regressions based on equation 1 where the dependent variable is the number of nurseries and Kitas per 1,000 children in the age group 0–3, $Z_{ly}$. The relevant sample in Column (1) corresponds to all households where the mothers were at least 22 years old at the time that the child was born, regardless of the eligibility of the mother to take a job-protected leave. Column (2) limits the analysis to all households in which mothers were eligible to take a job-protected leave. This corresponds to the main sample of analysis throughout Section 4. Finally, the sample in column (3) is restricted to the selected sample of households where the mothers reported to take a short maternal leave (a leave of at most 14 months). All specifications are estimated via OLS and include locality and birth cohort fixed effects. All standard errors are clustered at the locality level.
Table 7: Parameter estimates of the MTE curve

<table>
<thead>
<tr>
<th></th>
<th>1(Work)</th>
<th>Hours</th>
<th>Income</th>
<th>1(Welfare)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_{x,0}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I(Mother passed Abitur)</td>
<td>0.141</td>
<td>1.751</td>
<td>446.555</td>
<td>0.017</td>
</tr>
<tr>
<td></td>
<td>(0.081)</td>
<td>(2.223)</td>
<td>(187.697)</td>
<td>(0.040)</td>
</tr>
<tr>
<td>I(Foreign mother – Eastern Europe)</td>
<td>-0.123</td>
<td>-1.433</td>
<td>-69.939</td>
<td>0.170</td>
</tr>
<tr>
<td></td>
<td>(0.089)</td>
<td>(2.915)</td>
<td>(142.828)</td>
<td>(0.076)</td>
</tr>
<tr>
<td>I(Foreign mother - No Eastern Europe)</td>
<td>-0.090</td>
<td>0.553</td>
<td>-34.180</td>
<td>0.191</td>
</tr>
<tr>
<td></td>
<td>(0.084)</td>
<td>(2.493)</td>
<td>(164.179)</td>
<td>(0.084)</td>
</tr>
<tr>
<td>Age of mother at birth (in years)</td>
<td>-0.002</td>
<td>0.044</td>
<td>-13.743</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.148)</td>
<td>(10.342)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>Age of kid at interview (in years)</td>
<td>0.068</td>
<td>2.061</td>
<td>145.454</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>(0.017)</td>
<td>(0.560)</td>
<td>(36.619)</td>
<td>(0.012)</td>
</tr>
</tbody>
</table>

$\beta_x = \beta_{x,1} - \beta_{x,0}$

|                               |         |        |         |            |
| I(Mother passed Abitur)       | -0.215  | -2.975 | -191.891| -0.100     |
|                               | (0.114) | (3.371)| (259.853)| (0.057)    |
| I(Foreign mother – Eastern Europe) | 0.147  | 3.945  | -8.754  | -0.211     |
|                               | (0.156) | (5.126)| (252.424)| (0.128)    |
| I(Foreign mother - No Eastern Europe) | 0.023  | -2.007 | -132.112| -0.184     |
|                               | (0.124) | (3.952)| (265.807)| (0.121)    |
| Age of mother at birth (in years) | 0.028  | 0.533  | 94.573  | -0.020     |
|                               | (0.007) | (0.232)| (16.905)| (0.005)    |
| Age of kid at interview (in years) | -0.019 | -0.627 | -11.222 | 0.008      |
|                               | (0.019) | (0.605)| (38.685)| (0.012)    |

p-value of $H_0: \beta_x = 0$

|                               | 0.000   | 0.081  | 0.000   | 0.000      |

p-value of $H_0: \gamma_1 = \gamma_2 = 0$

|                               | 0.094   | 0.012  | 0.157   | 0.894      |

No. of observations

|                               | 12,445  | 12,445 | 11,676  | 12,094     |

Note: Table 7 presents the estimates of the MTE equation shown in equation (8). The $\beta_{x,0}$ coefficients refer to the effects of observable household characteristics on outcomes in the untreated state. The coefficients $\beta_x = \beta_{x,1} - \beta_{x,0}$ represent the gains from treatment, based on observable household characteristics. The first p-value corresponds to the p-value associated with the joint test that the two parameters of the second-order polynomial $k(p) = \gamma_1 p + \gamma_2 p^2$ are equal to zero. The second p-value is the p-value associated with the joint test that all the coefficients of $\beta_x$ are equal to 0. All standard errors are clustered at the locality level.
Table 8: Estimates of the rescaled ATE, ATT, ATUT, and LATE

<table>
<thead>
<tr>
<th></th>
<th>1(Work)</th>
<th>Hours worked</th>
<th>Monthly income</th>
<th>1(Receive transfers)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATE</td>
<td>0.214</td>
<td>14.687</td>
<td>779.543</td>
<td>-0.181</td>
</tr>
<tr>
<td></td>
<td>(0.199)</td>
<td>(7.053)</td>
<td>(502.802)</td>
<td>(0.141)</td>
</tr>
<tr>
<td>ATT</td>
<td>0.164</td>
<td>13.969</td>
<td>738.560</td>
<td>-0.167</td>
</tr>
<tr>
<td></td>
<td>(0.201)</td>
<td>(6.605)</td>
<td>(507.146)</td>
<td>(0.129)</td>
</tr>
<tr>
<td>ATUT</td>
<td>0.272</td>
<td>15.539</td>
<td>829.317</td>
<td>-0.192</td>
</tr>
<tr>
<td></td>
<td>(0.232)</td>
<td>(9.089)</td>
<td>(578.059)</td>
<td>(0.176)</td>
</tr>
<tr>
<td>LATE</td>
<td>0.100</td>
<td>10.688</td>
<td>699.12</td>
<td>0.017</td>
</tr>
<tr>
<td></td>
<td>(0.155)</td>
<td>(5.223)</td>
<td>(384.898)</td>
<td>(0.095)</td>
</tr>
<tr>
<td>No. of observations</td>
<td>12,445</td>
<td>12,445</td>
<td>11,676</td>
<td>12,094</td>
</tr>
</tbody>
</table>

Note: Table 8 presents rescaled estimates of the ATE, ATT, ATUT, and LATE. These estimates are derived from the estimated MTE curve, by integrating the MTE curve with the corresponding weights presented in Heckman and Vytlacil (1999). 95 percent confidence intervals are shown in parentheses.
Figure 1: Evolution of the number of daycare centers, seats, and children in childcare in Germany

Panel A: Nurseries and Kitas, per child aged 0–3

Panel B: Number of seats per child, ages 0–3 (upper bound)

Panel C: Percentage of children ages 0–3 in childcare

Panel D: Pedagogical staff per child in county, ages 0 – 7

Note: Figure 1 shows the evolution of alternative measures that capture different aspects about the availability of subsidized childcare across the 401 counties in Germany between 2007 and 2015. Panel A shows the evolution of the total number of nurseries and Kitas in the county per child in the age group 0–3. Panel B shows the evolution of the total number of local seats in childcare for children in ages 0–3, normalized by the number of children in that age group. This measure is an upper bound on the true measure of seats; see the main text for more details. Panel C instead shows the evolution of the proportion of children ages 0–3 in childcare. Finally, Panel D shows the evolution of the total number of pedagogical staff in the county, normalized by the number of children in the age group 0 – 7. In each of the graphs included in Panels A through Panel D, each box shows the 25th, 50th, and 75th percentiles (x_{25}, x_{50}, x_{75}) of the corresponding measure for the given year for the counties in West or East Germany. The whiskers correspond to the values x_{25} – 1.5 × (x_{75} – x_{25}) and x_{75} + 1.5 × (x_{75} – x_{25}).
Figure 2: Classification of German counties, by urbanicity level

Note: Figure 2 presents a map of the 401 local authorities at the county level in Germany. The borders of the 16 different states are highlighted in bold. The main geographical unit of analysis in the paper is defined as the intersection between state identifiers, the type of urbanicity of the county, the type of urbanicity of the region, and an indicator variable on whether the county has a municipality with 500,000 or more inhabitants. Based on this classification, counties are grouped into 89 different geographical units. The urban classification of counties and regions is based on the classification developed by the Federal Institute for Research on Building, Urban Affairs, and Spatial Development (BBSR) for the year 2016.
Figure 3: Cumulative density function of maternal leave duration

Panel A: West Germany

Panel B: East Germany

Note: Figure 3 presents the cumulative density function (cdf) of the total duration of maternal leave by aggregated birth cohort in the main sample. Panel A shows the cdfs of total duration of maternal leave for the different cohorts in West Germany, while Panel B shows the same cdfs for households in East Germany.
Figure 4: Relationship between maternal leave duration and normalized number of daycare centers in the locality

Panel A: Without household controls

Panel B: With household controls

Note: Figure 4 presents the relationship between the local number of nurseries and Kitas per child in the age group 0–3 and the maternal leave duration of households in the sample. Panel A shows a binned scatterplot of the residuals of both variables, after controlling for location and cohort fixed effects. Panel B shows the binned scatterplot of both variables, after additionally controlling for observable household characteristics. Histograms of the residualized versions of supply are plotted in the secondary axis.
Figure 5: Marginal effect of the local availability of childcare by year, relative to year of birth

Note: Figure 5 presents the marginal effect of the number of nurseries and Kitas per 1,000 children in the age group 0–3 at different periods of time, conditional on the number of daycare centers in the year of birth (when this variable should have an effect). The shaded gray region in the event study plots depict the 95 percent confidence interval for each year, relative to the year of birth. The models include the set of controls of the main specification shown in column (2) of Table 2. The coefficients on the local number of daycare centers per 1,000 children ages 0–3 for the years -3 to -1 and 1 to 3 are all conditional on the normalized number of daycare centers in year 0 (the corresponding birth cohort of the child). The coefficient for the local number of daycare centers per 1,000 children ages 0–3 in the year of birth is based on a model with no other periods included.
Figure 6: Binned scatterplot of the estimated propensity score as a function of the instrument, by maternal education

Note: Figure 6 shows the relationship between the number of nurseries and Kitas per 1,000 children ages 0–3 in the locality in the year of birth (Z) and the estimated propensity score that the household takes a short leave \( P(X, Z) \), separately by the type of school-leaving certificate of the mother. In this binned scatterplot, households are grouped into 20 different bins based on the value of \( Z \), separately for each education level. Then the average propensity score in each bin is calculated and shown in the y-axis. The solid lines show the fitted regression lines of the propensity scores \( P(X, Z) \) on \( Z \), separately by the level of maternal education.
Figure 7: Propensity score and common support of treated and non-treated households

Panel A: Without interactions between $Z$ and $X$

Panel B: Including interactions between $Z$ and $X$

Note: Figure 7 presents the estimated propensity scores of treatment take-up for treated and untreated households. Panel A shows the estimated propensity scores for households from a specification where the only instrument is the local availability of childcare. Instead, Panel B shows the estimated propensity scores for households from a specification where the interactions between the local availability of childcare and observed household characteristics are also included as instruments.
Figure 8: Estimated MTE curves

Panel A: 1 (Hours worked > 10)  Panel B: Total hours worked

Panel C: Monthly maternal income  Panel D: 1 (Household receives transfers)

Note: Panels A – D of Figure 8 show the estimated MTE curves as a function of $V$, the unobserved resistance to treatment. The MTE curve is estimated at the sample mean of the observed households $\bar{X}$ and is shown in red. The 95% confidence intervals of the MTE curve are shown in gray. Additionally, the dashed horizontal line shows the rescaled Average Treatment Effect (ATE) in the area of common support.
Appendix A: Nonparametric identification of the Marginal Treatment Effect

For the sake of completeness, I present a brief summary of the argument about the nonparametric identification of $MTE(x, v)$. Heckman and Vytlacil (1999) show that if the conditional independence assumption holds, then:

$$E[Y \mid X = x, P(D = 1 \mid X, Z) = p] = \mu_0(x) - p \times (\mu_1(x) - \mu_0(x)) + K(x, p)$$

where $K(x, p) = p \times E[U_1 - U_0 \mid X = x, V \leq p]$. The econometrician observes the empirical analog of the conditional expectation in the left-hand side in the data. After taking the derivative of both sides with respect to $p$ and evaluating this derivative at $p = v$, you obtain the following result:

$$\frac{\partial E[Y \mid X = x, P(D = 1 \mid X, Z) = p]}{\partial p} \bigg|_{p=v} = \mu_1(x) - \mu_0(x) + k(x, v) \equiv MTE(x, v)$$

where $k(x, v) = E[U_1 - U_0 \mid X = x, V = v]$. The empirical analog of the left-hand side of this expression is again observed by the econometrician, since the econometrician observes realizations of the joint distribution of $(Y, X, Z, D)$.

Thus, if Assumption 1 holds, the $MTE(x, v)$ is nonparametrically identified over the support of $P(X, Z)$, conditional on $X = x$. The identification argument is also constructive, as it shows how to estimate $MTE(x, v)$ by taking the derivative of the conditional expectation of the observed outcome with respect to the propensity score $p$, holding the value of $x$ fixed.
Appendix Tables and Figures

Table A-1: Effects of local childcare availability on maternal leave duration, West Germany only

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Childcare centers (per 1,000 children)</td>
<td>-0.347</td>
<td>-0.358</td>
<td>-0.320</td>
<td>-0.366</td>
</tr>
<tr>
<td>1(Mother passed Abitur)</td>
<td>-3.563</td>
<td>-3.562</td>
<td>-3.565</td>
<td></td>
</tr>
<tr>
<td>1(Foreign mother, Eastern Europe)</td>
<td>0.298</td>
<td>0.302</td>
<td>0.296</td>
<td></td>
</tr>
<tr>
<td>1(Foreign mother, not Eastern Europe)</td>
<td>-0.460</td>
<td>-0.458</td>
<td>-0.467</td>
<td></td>
</tr>
<tr>
<td>Age of mother at birth</td>
<td>0.152</td>
<td>0.152</td>
<td>0.152</td>
<td></td>
</tr>
<tr>
<td>Local unemployment rate</td>
<td>0.813</td>
<td>0.713</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local number of births per women ages 15–45</td>
<td>-0.009</td>
<td>0.038</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proxy for childcare quality</td>
<td>4.854</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Locality FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Birth cohort FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Mean leave duration (in months)</td>
<td>18.80</td>
<td>18.80</td>
<td>18.80</td>
<td>18.80</td>
</tr>
<tr>
<td>F-statistic of $H_0 : \beta = 0$</td>
<td>5.782</td>
<td>6.116</td>
<td>5.537</td>
<td>8.268</td>
</tr>
<tr>
<td>p-value of $H_0 : \beta = 0$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Number of observations</td>
<td>7,731</td>
<td>7,731</td>
<td>7,731</td>
<td>7,731</td>
</tr>
</tbody>
</table>

Note: Appendix Table A-1 presents the estimated coefficients of the generalized difference-in-differences shown in Equation 1 for the sub-sample of households in West Germany. All regressions are estimated by Ordinary Least Squares. The dependent variable is the maternal leave duration in months. The main independent variable is the number of nurseries and Kits per 1,000 children in age group 0–3 in the locality. All specifications include locality and birth cohort fixed effects. The specification in column (2) includes observed household characteristics. The specification in column (3) additionally controls for the local unemployment rate and number of births per women in the age group 15–45. Finally, column (4) also includes an index of childcare quality at the local level. All standard errors are clustered at the locality level.
Table A-2: Marginal effects of local childcare availability on take-up of short leaves, selection equation specification

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of nurseries and Kitas</td>
<td>0.018</td>
<td>0.020</td>
<td>0.021</td>
<td>0.020</td>
<td>0.018</td>
</tr>
<tr>
<td>(per 1,000 children ages 0–3)</td>
<td>(0.006)</td>
<td>(0.007)</td>
<td>(0.007)</td>
<td>(0.007)</td>
<td>(0.006)</td>
</tr>
<tr>
<td>Locality FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Birth cohort FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Mean dependent variable</td>
<td>0.637</td>
<td>0.637</td>
<td>0.637</td>
<td>0.637</td>
<td>0.637</td>
</tr>
<tr>
<td>p-value of $H_0 : \beta = 0$</td>
<td>0.004</td>
<td>0.002</td>
<td>0.001</td>
<td>0.006</td>
<td>0.004</td>
</tr>
<tr>
<td>Number of observations</td>
<td>12,642</td>
<td>12,642</td>
<td>12,642</td>
<td>12,642</td>
<td>12,642</td>
</tr>
</tbody>
</table>

Note: Table A-2 presents the estimated average marginal effect of a change in the number of nurseries and Kitas per 1,000 children in the age group 0–3 on the probability that mothers take short leaves based on equation (5). All columns have been estimated using probit. All specifications include observed household characteristics, as well as locality and birth cohort fixed effects. Column (1) is the baseline model. Column (2) includes the local unemployment rate and the number of births per women ages 15–45 in the locality during the year of birth. Column (3) additionally includes the index of childcare quality at the local level in the year of birth of the child. Column (4) additionally includes two time-varying characteristics at the local level in the year when the child turns 3 years old: the proportion of children in the age range 3–6 that attends kindergarten in a full-time manner and an indicator variable on whether kindergarten is free. Finally, column (5) has the same controls as column (1) and also the number of children in the household at the time of the interview. All standard errors are clustered at the locality level.
Table A-3: Sensitivity analysis of the ATE of short leaves on maternal hours worked per week

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(6.186)</td>
<td>(4.786)</td>
<td>(4.832)</td>
<td>(4.837)</td>
<td>(5.123)</td>
</tr>
<tr>
<td></td>
<td>(6.398)</td>
<td>(5.527)</td>
<td>(5.310)</td>
<td>(5.310)</td>
<td>(5.316)</td>
</tr>
<tr>
<td>ATUT</td>
<td>15.539</td>
<td>15.313</td>
<td>14.330</td>
<td>10.246</td>
<td>10.379</td>
</tr>
<tr>
<td>Number of observations</td>
<td>12,445</td>
<td>12,445</td>
<td>12,445</td>
<td>12,445</td>
<td>12,445</td>
</tr>
</tbody>
</table>

Note: Table A-3 presents a sensitivity analysis of the ATE of short leaves (relative to long leaves) on the maternal number of hours worked per week. Column (1) presents the baseline MTE model. Column (2) includes the local unemployment rate and the number of births per women ages 15–45 in the locality during the year of birth in the selection equation and potential outcomes equations. Column (3) additionally includes the index of childcare quality at the local level in the year of birth of the child in the selection equation and potential outcomes equations. Column (4) additionally includes two time-varying characteristics at the local level in the year when the child turns 3 years old: the proportion of children in the age range 3–6 that attends kindergarten in a full-time manner and an indicator variable on whether kindergarten is free in the selection equation and potential outcomes equations. Finally, column (5) has the same controls as column (1) but also the number of children in the household at the time of the interview in the selection equation and potential outcomes equations. All standard errors are clustered at the locality level.
Figure A-1: Relationship between maternal leave duration and alternative measures of local availability of childcare

Panel A: Number of seats per child in age group 0–3 (upper bound)

Panel B: Proportion of pedagogical staff per child, 0 - 7

Panel C: Proportion of children 0–3 in a childcare center

Note: Figure A-1 presents the relationship between different measures of local availability of subsidized childcare for children ages 0–3 and maternal leave duration in the sample. For the explanation of the different measures see the main body of text. The graphs in the left-hand side show a binned scatterplot of the residuals of the duration of leave and the availability measure, after controlling for location and cohort fixed effects. The graphs in the right additionally control for observable household characteristics. The probability associated with the histograms of the residualized measure of supply is plotted in the secondary axis.
Figure A-2: Sensitivity of estimated marginal effects of changes in local availability of childcare to different monthly cut-offs to define “short” leaves

Note: Figure A-2 shows the estimated marginal effects of an increase in the number of daycare centers per 1,000 children on the age group 0–3 on the probability that mothers take a “short” leave using different cut-offs to define “short” leaves. All the estimated marginal effects come from probit specifications that also include observable household characteristics, and location and birth cohort fixed effects as exogenous regressors. The 95 percent confidence intervals of the estimated marginal effects are also shown in Figure A-2. On the y-axis, the $\chi^2$ value of the hypothesis test that the marginal effect is equal to zero for each of the different cut-offs is also plotted in the secondary y-axis. All standard errors are clustered at the locality level.