

Gambaro, Ludovica; Marcus, Jan; Peter, Frauke

Article — Accepted Manuscript (Postprint)

School entry, afternoon care, and mothers' labour supply

Empirical Economics

Provided in Cooperation with:

German Institute for Economic Research (DIW Berlin)

Suggested Citation: Gambaro, Ludovica; Marcus, Jan; Peter, Frauke (2019) : School entry, afternoon care, and mothers' labour supply, Empirical Economics, Springer, Berlin, Vol. 57, Iss. 3, pp. 769-803, <https://doi.org/10.1007/s00181-018-1462-3>

This Version is available at:

<https://hdl.handle.net/10419/214192>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

School entry, afternoon care and mothers' labour supply

Ludovica Gambaro

DIW Berlin and UCL Institute of Education

Jan Marcus*

Universität Hamburg and DIW Berlin

Frauke Peter

DIW Berlin

Abstract

The availability of childcare is a crucial factor for mothers' labour force participation. While most of the literature examines childcare for preschool children, we specifically focus on primary school aged children, estimating the effect of formal afternoon care on maternal labour supply. To do so, we use a novel matching technique, entropy balancing, and draw on the rich and longitudinal data of the German Socio-Economic Panel (SOEP). We show that children's afternoon care increases mothers' employment rates and their working hours. To confirm the robustness of our results, we conduct a series of sensitivity analyses and apply a newly proposed method to assess possible bias from omitted variables. Our findings highlight how childcare availability shapes maternal employment patterns well after school entry.

Keywords: Afternoon care, Maternal labour supply, All-day schools, Entropy balancing
JEL: J13, J63, J65

Acknowledgements

We gratefully acknowledge support by the College for Interdisciplinary Educational Research (CIDER). Moreover, we thank C. Katharina Spieß, Adam Lederer, Janina Nemitz, seminar participants at the University of Chicago and DIW Berlin, as well as participants of the GEBF 2016 and the 2016 ESPE conferences for valuable comments.

*Corresponding author

Email addresses: lgambaro@diw.de (Ludovica Gambaro), Jan.Marcus@uni-hamburg.de (Jan Marcus), fpeter@diw.de (Frauke Peter)

1. Introduction

Women’s labour force participation has increased dramatically, significantly narrowing the gender employment gap. Yet women with children remain less likely to work in the labour market than other women or men, and when they do work, they tend to work fewer hours. Such lesser engagement has important consequences from both social and individual perspectives. In many advanced economies, the weak attachment of mothers to the labour market leads to a systematic underutilisation of their human capital, often exacerbating an already unfavourable employment/population ratio. At the same time, the economic position of mothers is affected both in the short- and long-term, as employment interruptions and lower working hours not only result in an immediate loss of earnings, but also tend to place them on a permanently lower earnings trajectory (e.g. Joshi et al., 1996; Sigle-Rushton and Waldfogel, 2007; Waldfogel, 1998).

Family policies appear to be successful in reducing the cost of motherhood. Indirect evidence comes from cross-country comparisons, with mothers maintaining stronger ties to the labour market in places where governmental support is more extensive (Gornick and Meyers, 2003; OECD, 2011). There is also persuasive evidence on the effectiveness of specific family policies in individual countries, as a large number of studies examine the effect of an individual policy change on maternal employment. This body of research considers different policy reforms, including leave policies (e.g. Berger et al., 2005; Nollenberger and Rodríguez-Planas, 2015), childcare subsidies (e.g. Baker et al., 2008; Brilli et al., 2013; Lefevbre and Merrigan, 2008; Schober and Spiess, 2015), and public pre-school education (e.g. Berlinski and Galiani, 2007; Blau and Currie, 2006; Cascio, 2009; Fitzpatrick, 2012; Goux and Maurin, 2010). However, it focuses almost exclusively on children under the compulsory school age.

Yet caring responsibilities for children do not end once they enter primary school. Although schools effectively provide what is free and universal childcare, maternal labour supply patterns continue to be influenced by the presence of children, even when they are of school age. Not only does school life impose a new set of demands on parental time, but, crucially, in most countries regular school hours are part-time and not compatible with a full-time working week (OECD, 2011; Plantenga and Remery, 2017). In most countries, primary school children start the school day between 8:00 am and 9:00 am and finish between 1:00 pm and 3:00 pm; but hours are not always evenly spread across the week. For example, until 2014, schools in France operated a four-days a week schedule, while most schools in Belgium and the Netherlands are not open on Wednesday afternoon (Plantenga and Remery, 2017). While a few countries, including Sweden and Denmark,

have a well-organised infrastructure of out-of-school services tailored on school opening hours, in most countries provision is patchy. The problem of organising afternoon care is especially acute in countries where the school day is traditionally limited to only the morning and where children typically go home at lunch time, including, for example, Austria, Chile, Germany, Mexico, and Switzerland (Allemann-Ghionda, 2009; OECD, 2011). Most of these countries are implementing reforms to increase the length of time children spend in school or increase the availability of after-school programmes (OECD, 2015).

In this paper, we examine the case of Germany, where, since the early 2000s, policy makers have sought to support maternal employment by extending the time school children spend in formal afternoon care (Marcus et al., 2016). As a result, Germany has moved from an exclusively half-day school system to one where afternoon care is increasingly available, either because schools operate a full-day schedule or because after-school programmes, often based in school facilities, offer additional activities. The change amounts to an extension of the public school system, as schools remain the cornerstone of this expanded care provision, even when they do not provide the service themselves. Germany is an interesting case for a number of reasons. Unlike the Nordic countries or France, (West) Germany has long been characterised by low maternal employment: not only there is a large employment gap between mothers and childless women, but among working mothers short part-time work (less than 20 hours a week) is the dominant employment arrangement (Daly, 2000; Knittel et al., 2014; Lewis et al., 2008). Informal childcare arrangements are not common, as they are in Southern European countries such as Italy (Bettio and Plantenga, 2004). Further, a highly regulated labour market makes it difficult to meet the demand for afternoon care: unlike the US or the UK, affordable private services staffed by a low-paid and low-qualified workforce is not available (Morgan, 2005). At the same time, Germany has witnessed a radical policy shift with the development of a number of family friendly policies aimed at easing the reconciliation between family responsibilities and employment (Bauernschuster and Schlotter, 2015; Schober and Spiess, 2015). These changes have occurred against the background of an expanding economy that was not hit by the great recession to the same extent as other countries. So, from a public policy perspective, Germany could be thought of as “low hanging fruit”, a context in which we would expect the availability of childcare to significantly influence maternal labour supply.

Of the considerable literature on childcare and maternal employment, studies evaluating the impact of pre-school provision one or two years before compulsory school

age are most closely related to our paper. The 1970s roll-out of kindergarten programmes for five year olds in the US is a prominent example. Gelbach (2002) finds that enrolment increases labour market participation of both married and single mothers, albeit, for the latter group, only when they do not have an additional younger child. More recent results from Cascio (2009) and Fitzpatrick (2012) are mixed, as they show that kindergarten increases the probability of working among single mothers with one child, but with no effect on single mothers with an additional younger child or married mothers. There are studies from other countries that also exploit reforms expanding preschool education throughout the 1990s. For example, Berlinski and Galiani (2007) show how the construction of pre-primary school facilities in Argentina helped raise enrolment among 3-5 year olds and, in turn, maternal employment. Nollenberger and Rodríguez-Planas (2015) estimate that, notwithstanding low labour demand, the fivefold expansion of universal preschool education for three year olds in Spain increased maternal employment by almost 10%. Bauernschuster and Schlotter (2015) examine the impact of the increase in Kindergarten attendance by 3 year olds and older in Germany. Their results indicate a large and positive effect on maternal labour force participation. In all these countries where reforms occurred in the 1990s, maternal employment at the time of the change in preschool availability was rather low and, in the case of Germany, even among highly educated mothers living in strong labour markets. By contrast, in Scandinavian countries, public childcare is found to have no impact on the already high maternal labour force participation rate (e.g. Havnes and Mogstad, 2011; Kosonen, 2014; Lundin et al., 2008; Simonsen, 2010).

There are only a few studies specifically examining the effect of reforms extending school opening hours or after-school care programmes. Evidence from Chile, where the school day was increased by two hours at the end of the 1990s, indicates a positive effect on female employment (Berthelon et al., 2015). Felfe et al. (2016) examine the case of Switzerland, where the legal right to an after-school care place was introduced by different cantons in different years. They find a positive effect on full-time employment among mothers, but no effect on the employment rate. Paternal employment, instead, did not appear to respond to the increase in after-school care. There is also an emerging literature on the German case, focusing on a specific federal policy programme launched in 2003, under which schools are expected to provide children with lunch and afternoon care following the regular morning instruction hours. One of the policy objectives of the programme is to increase maternal labour participation. Using a structural micro-simulation, Beblo et al. (2005) indeed show that such a programme would increase

maternal labour force participation. However, it is proving more challenging to estimate the impact of the policy using standard quasi-experimental evaluation methods. While there are three working papers (Dehos and Paul, 2017; Nemitz, 2015; Shure, 2016) attempting to exploit the staggered implementation of this programme as a source of exogenous variation, accounts of the actual roll out suggest that implementation was neither random (Wiezorek et al., 2011) nor isolated from childcare policy at the local level (Autorengruppe Bildungsberichterstattung, 2016; Lange, 2016). As we explain more fully in section 2, this single federal programme contributes to the expansion of afternoon care, but in a way that reflects an intricate pattern of different local labour market conditions, political priorities, and existing afternoon care services. Without data on all these other contextual factors, it is unlikely that the variations in the roll-out of this specific programme are unrelated to maternal employment.

In this paper, we take an approach that is more suitable to examine the kind of expansion in afternoon care that occurred in Germany. Drawing on the German Socio-Economic Panel (SOEP), we investigate the employment patterns of mothers in the year that their child enters compulsory schooling and the year before. We examine two different outcomes: i) being in employment; and ii) actual hours worked per week, thus distinguishing between the extensive and intensive margin. We compare mothers whose child only attends school in the morning to similar mothers whose child also attends an afternoon care service. In order to make the two groups more similar, we apply a rather novel non-parametric matching estimator, entropy balancing (see Hainmueller, 2012). The matching procedure has several advantages compared to common propensity score methods and makes use of a rich set of information about the mothers, their children, and their partners in order to make our strategy robust against selection on observables. Further, by considering the lagged value of the outcome variable, our empirical strategy takes into account selection on unobserved variables that do not change over a short period of time (such as attitudes toward work and family). We address some concerns about reverse causality by exploiting a specific feature of enrolment procedures, whereby children receiving afternoon care organised under the auspices of the school can only register at the beginning of the school year. Despite the fact that we employ more than 100 conditioning variables and that these contain the mother’s detailed labour market history as well as some usually unobserved variables like desired working hours and job search behaviours, omitted variables might still be present. Therefore, we apply the method proposed by Oster (2013, 2016) to assess the robustness of our results to omitted variable bias. This method exploits the fact that the bias from observed variables is informative

of the bias from omitted variables, assuming that there is some kind of proportionality between the two biases.

We find that a child’s participation in afternoon care during the first year of primary school increases their mother’s employment. By taking into account the different employment patterns prior to school entry, we further show that a mother who did not work before is more likely to take up paid work (+11.4 percentage points), while among mothers who already worked prior their child’s school entry, afternoon care leads to an average increase in working hours by about 2.6 hours per week. There is little evidence that this increase in maternal employment crowds out paternal employment or other childcare arrangements. Our results are robust to various sets of control variables, different sample restrictions, and alternative estimations techniques. Further, the Oster (2013, 2016) method suggests that the impact of omitted variables must be substantively stronger than that of the included control variables in order to completely explain the effects of afternoon care.

The remainder of the paper is organised as follows. Section 2 briefly describes the institutional context. In Section 3 we present our empirical strategy and Section 4 provides details about the data used. Section 5 discusses our results and Section 6 comprises various robustness checks before Section 7 concludes.

2. Institutional context

Afternoon care

In Germany children enter formal schooling when they are six years old and attend primary schools for four years before transitioning into secondary education.¹ Primary schools have traditionally been organised on a half-day basis, with lessons taking place only in the morning from 8:00 am and children returning home for lunch by 1:00 pm. In the German Democratic Republic (GDR; East Germany) this system was progressively supplemented with after-school programmes provided in the school building (*Schulhorte*) but organised under different auspices (Mattes, 2011). In West Germany, instead, the traditional half-day structure of the school system remained in place until the early 2000s and relied on the presence of mothers at home (Hagemann, 2006).

The lack of afternoon care, with its consequences for maternal employment and possible contribution to educational inequalities, gained great political prominence only

¹In the states of Berlin, Brandenburg, and Mecklenburg-Vorpommern, the transition to secondary school occurs after grade 6.

in the early 2000s and culminated in the launch of a federal government flagship programme, “The Future of Education and Care” (*Investitionsprogramm Zukunft Bildung und Betreuung* or *IZBB*). This policy initiative provides funds designated to support both primary and secondary schools remain open in the afternoon, offering after-school activities. Schools participating in the programme are now known as “all-day schools”. The name is somewhat misleading, as among primary schools the large majority of “all-day schools” do not have instruction hours in the afternoon. Instead, they offer optional extra-curricular activities after normal lessons. Despite IZBB being a single high profile initiative, each federal state administers the funds differently. This is in part due to German governance structure, whereby education policy is the exclusive responsibility of each individual federal state. For example, some states concentrate their funding on secondary schools, while others opt to support primary schools or give priority to working parents. In many states, funds within the IZBB framework are distributed on a “first-come-first-serve” rule among applying schools. This means that expectations of parental demand for afternoon care in the school catchment area are likely to influence the roll-out of the programme. Because of this complex funding and implementation structure, IZBB-sponsored afternoon programmes have evolved differently across the country, reflecting different political priorities and resources. In particular, differential expectations or policy aims in relation to maternal employment may play a role in how federal funds are spent. As such, it is difficult to exploit these differences in development rates as a source of exogenous variation without data capturing how local level policies and individual school practices mediate access to this specific federal initiative (Wiezorek et al., 2011).

Notwithstanding these variations, all-day schools across the whole of Germany have been increasing, with over 53% of all primary schools offering afternoon care in 2014 (KMK, 2016). In practice, this means that children attending these schools have the option of remaining at school, eating lunch and taking part in extracurricular activities. The activities are offered either directly by the school, albeit not by teaching staff, or by external providers operating on school premises. Afternoon programmes vary greatly in their content, with some closely linked to the morning school lessons, while others offer other types of activities, such as sport or music. Enrolment in the afternoon programme is voluntary in almost all primary schools (Marcus et al., 2013),² with enrolment only taking place at the beginning of the school year (Federal Ministry of Education and Research,

²Less than 5% of primary schools are all-day schools in the strict sense of operating an all-day schedule for all the children enrolled (Marcus et al., 2013).

2016). This means that parents do not have the flexibility to adjust their children’s attendance on the basis of job opportunities that may arise during the school year.

Afternoon services can also be provided independently from primary schools, resulting in a combination of primary schooling in the morning and an after-school programme in the afternoon (*Hort*). Within this type of programme, children are picked up from school, given lunch, and offered a variety of activities. This is most commonly offered at community centres, but sometimes on school premises. Like all-day schools, this type of programme can operate in very different ways, with looser or tighter links to primary schools. While the precise educational and pedagogical content of the different programmes are likely to matter to children’s development, here we intentionally leave aside these aspects and focus instead on those organisational features that are most relevant to enabling mothers to work. From this perspective, there is considerable overlap between the two types of programmes. In both cases, children are provided with lunch and spend their afternoon hours in a supervised environment with learning and enrichment opportunities. Afternoon activities take place four or five days a week, usually lasting until 3:00 or 4:00 pm, depending on the programme (Holtappels et al., 2008), providing a considerable extension of childcare coverage relative to morning only school attendance.

Perhaps it is not surprising that policy makers see these two types of programmes as either substitutes or, essentially, equal. Thus, with the role out of the IZBB programme some Western states (such as Berlin, Hamburg, North Rhine-Westphalia) have substituted *Hort* programmes with all-day schools, while other states have fostered the increase of both *Hort* programmes and all-day schools. In the remainder of this paper, we use the terms “afternoon care” to refer to the two formal services described above. When we need to distinguish between specific types of programmes, we refer to afternoon care under the auspice of the school as all-day schools (*Ganztagsschule*) and to all other formal afternoon care as after-school programmes (*Hort*). We exclude other forms of care that may be provided by friends or relatives as well as other privately arranged out-of-school activities.

Maternal labour supply

In Germany, the employment rates of mothers differ greatly between West and East. In particular, in West Germany, where in 2012 82% of all mothers with dependent children lived, their labour force participation has historically been low. However, maternal labour supply is growing considerably, increasing from 59% in 2000 to 66% in 2012 among mothers with children younger than 18 (Knittel et al., 2014). Between 2006 and 2012, increases were particularly pronounced among mothers with children around school

entry age: the employment rates of mothers with children aged 4-6 and 6-8 rose by 8.3 and 8.8 percentage points, respectively (Knittel et al., 2014). Although mothers are increasingly in employment, they mainly work part-time, with more than three-quarters of employed mothers working less than 32 hours a week in 2012. This pattern does not vary substantially with the age of the youngest child. In 2012, working less than 32 hours a week was as common among mothers with children aged between 4 and 6 as among mothers of children aged between 6 and 8 (about 74%) (Knittel et al., 2014). Mothers increase their hours as their children grow, but compulsory school entry does not mark an abrupt increase in maternal labour market engagement. An increase in full-time work among mothers appears once children are 10 or older and, thus, more capable of self-care. The start of compulsory schooling at age 6 may even lead to a reduction in maternal employment or working hours in areas where the increasing availability of full time early childhood education and care places has not been matched by an extension of childcare coverage for primary school children. For example, while early childhood education and care centres have traditionally only been open in the morning in West Germany, in 2015 39% of children aged three to six had a full day place (Autorengruppe Bildungsberichterstattung, 2016). Overall, working part-time appears to be the way German mothers reconcile paid work with caring responsibilities for young children, yet Wunder and Heineck (2013) show that a substantial share of mothers (26%) would like to work longer hours.

Within such a context, afternoon care can be a powerful policy lever to support maternal employment. First, it can help mothers enter the labour market, giving them greater flexibility to take jobs that do not closely match half-day schools' opening hours. Second, afternoon care can help mothers already in employment extend their working hours, either within a part-time working arrangement or by moving to full-time. We test these hypotheses in what follows.

3. Empirical strategy

Our empirical strategy seeks to identify the impact of afternoon care for primary school children on maternal labour force participation. We investigate the role of both all-day schools (*Ganztagsschule*) and after-school programmes organised by other providers (*Hort*). Both types provide subsidized public afternoon care and, in practice, cannot always be distinguished from one another (Lange, 2015). Our identification strategy relies on a matching approach that makes also use of the pre-treatment outcomes. The longitudinal nature of our data allows controlling for mothers' pre-treatment labour force

participation while matching is performed by entropy balancing. The general idea of the estimation strategy is straightforward. We examine maternal employment patterns before and after their child enters school, comparing those whose child receives afternoon care (treatment group) to those whose child only attends school in the morning (control group). In order to make the control group children similar to the treatment group, we apply the non-parametric entropy balancing (EB).³ This technique reweights the observations in the control group in such a way that they have the same mean and variance for all included variables as the treatment group.⁴ We opt for entropy balancing over the more conventional propensity methods for a number of reasons. First, EB is more effective at reducing the imbalance between treatment and control group characteristics and, unlike propensity score methods, never produces a worse balance. Second, while the covariate balance is only checked for the means in most propensity score applications, EB allows for balancing both the mean and variance of each individual variable, thus further enhancing the balance between the two groups. Third, EB is fully non-parametric and does not rely on functional form assumptions necessary for the propensity score equation. Fourth, entropy balancing spares the burdensome iterations of propensity score methods between estimating the propensity score, checking for covariate balance and readjusting the propensity score model to achieve a better balance.

However, as in all propensity score methods, entropy balancing requires the inclusion of all variables that simultaneously affect the probability of children’s participation in

³We use the user-written programme “ebalance” (Hainmueller and Xu, 2013) in Stata to implement entropy balancing.

⁴More specifically, entropy balancing solves equations of the following form for ω_1 - ω_n :

$$\begin{aligned}\bar{x}_1^T &= \omega_1 \cdot x_{11} + \omega_2 \cdot x_{12} + \dots + \omega_n \cdot x_{1n} \\ \bar{x}_2^T &= \omega_1 \cdot x_{21} + \omega_2 \cdot x_{22} + \dots + \omega_n \cdot x_{2n} \\ &\vdots \\ \bar{x}_j^T &= \omega_1 \cdot x_{j1} + \omega_2 \cdot x_{j2} + \dots + \omega_n \cdot x_{jn}\end{aligned}$$

\bar{x}_j^T denotes the mean of variable x_j in the treatment group, x_{jn} the value of the n -th control group observation in the j -th conditioning variable and ω_n denotes the weight of the n -th control group member. Note that while this discussion focuses on the first moments (means), the argument for the second moment (variance) follows analogously. Entropy balancing imposes the restriction that all ω_n must be non-negative (i.e. no observation receives a negative weight). While there are generally more control group observations than restrictions, i.e., $N > J$, usually more than one weighting scheme ω_1 - ω_n solves this set of equations. Out of the many possible weighting schemes that fulfil these conditions, entropy balancing selects the weighting scheme in which the weights deviate as little as possible from equal weights - where distance is measured by the eponymous entropy divergence (Kullback, 1959).

afternoon care and maternal employment (conditional independence assumption, CIA). Because we work with longitudinal data and observe mothers both the year their child starts school and the year before, we can take several steps to make it more likely for the CIA to hold. We start by including the employment status of mothers when the child is below compulsory school age; this is the “pre-treatment outcome” component of our model. In particular, we specify whether mothers work at all and the number of actual hours they work. We include the same information on their labour force participation from *two* years before the child’s school entry. This way we compare mothers with similar employment trajectories. We also pay particular attention to work plans and motivation, including information that is often unobserved in other datasets, such as the number of desired working hours, job search behaviour, and working intentions for those not employed. In addition, we include information on children’s childcare attendance the year before school entry. We distinguish between institutional care and informal care, thus capturing preferences for centre-based care and whether relatives are available. Family socio-economic and demographic characteristics, such as education levels, income, and family composition, are also accounted for by a large set of variables. Finally, we add regional indicators to capture disparities in economic conditions. In short, our matching strategy relies on an extensive set of observables that go a long way to capture work and childcare preferences to the fullest extent possible.

Our set of control variables consists of factors that might affect the treatment and factors that might affect the outcome. Note that the conditional independence assumption states that the estimator is biased only if a variable that is related to both treatment and outcome is not included. Therefore, a failure to control for any variable in the two sets of factors only results in a biased estimator if the omitted variable should be included in both sets of factors. The matching procedure makes our strategy robust against selection on observables. Further, by considering the lagged dependent variable, our empirical strategy also takes into account selection on unobservable characteristics that are likely to remain stable over time, such as work-family attitudes. We include all control variables, not only in the entropy balancing step, but also in the regression equation. This makes the estimator double-robust (Bang and Robins, 2005) and also increases the precision of the estimates as the control variables reduce the unexplained variance in the outcome. Hence, we estimate the average treatment effect on the treated (ATT) using the following equation:

$$ATT = \sum_{k \in T} \left[(Y_{1k} - X'_k \hat{\beta}) - \sum_{l \in C} W_{k,l}(X) (Y_{0l} - X'_l \hat{\beta}) \right], \quad (1)$$

where Y_{1k} and Y_{0l} denote the observed outcomes of individuals in the treatment (T) and control group (C), respectively. X depicts the vector of control variables including the lagged outcome, while $W_{k,l}(X)$ refers to the weights from entropy balancing and, hence, depends on X . $\hat{\beta}$ denotes the vector of estimated coefficients from the weighted regression of Y on all control variables. Eq. (1) shows that the control variables are used both for entropy balancing and regression-adjustment.⁵ All reported standard errors are robust to heteroscedasticity and clustered at the mother level.

In practice, our empirical strategy seeks to ensure that variation in afternoon care attendance does not stem from mothers' characteristics, their propensity to work, or their attitudes towards childcare. Instead, when comparing mothers who differ only by their children's attendance in afternoon care, our assumption is that such variation is only due to the availability of formal afternoon care. We expect a child's attendance in afternoon care would allow their mother to search and take up jobs, while, by the same token, absence of afternoon care would lead a mother to turn down a job offer because of her care commitment. The strength of our approach is that our rich set of conditioning variables allows controlling directly for factors related to the demand for afternoon care, leaving variation in availability to be driven by exogenous idiosyncratic factors.⁶ Note that using the staggered expansion of afternoon care as an instrumental variable for a child's attendance is likely to incur into the problem that expansion itself was guided by demand at the very local level, especially since schools themselves have considerable autonomy on whether to offer the service or not.

There are two main threats to our identification strategy: omitted variables and reverse causality. While our identification strategy includes several features against these two threats, in the robustness section we apply the method developed by Oster (2016) to assess robustness to bias from those unobserved variables that may have a time-varying impact on maternal employment and for which we cannot control. We further address some concerns of reverse causality by exploiting a specific feature related to the timing of maternal employment choices and enrolment in afternoon care.

⁵As the treatment indicator is almost orthogonal to the control variables after entropy balancing (treatment and weighted control group have the same means in all control variables), the inclusion of control variables in the regression step does not change the estimated treatment effect substantially but rather increases its precision.

⁶More specifically, we not only control for the economic situation in the mother's region, for state fixed effects, and for the degree of urbanisation, but we also include variables that capture preferences for childcare policies (e.g., the child's care arrangements before school entry) and maternal work preferences (e.g., maternal work intentions and labour market history).

4. Data

The analyses in this study are based on the German Socio-Economic Panel (SOEP). The SOEP is an annual nationwide random panel survey of German households, carried out since 1984 (see Wagner et al., 2007). Currently, it covers more than 30,000 individuals in approximately 17,000 households. SOEP has several advantages for the present analysis. First, it is among the few nationally representative datasets in Germany that includes information on participation in afternoon care services. Second, the longitudinal design allows comparing maternal labour force participation before and after their child’s school start. Third, the dataset is especially rich, with detailed information on children and their childcare usage, on parents and their employment histories, as well as on the entire household. Fourth, SOEP comprises information that is usually unobserved in other data, such as job search behaviour and the intention to work for non-working individuals. Finally, the data include the date of the interview and detailed calendar information on individual labour force statuses, which helps to mitigate concerns that mothers’ decisions to work precede the children’s enrolment in afternoon care services.

4.1. Sample selection

Our focus is on the change in mothers’ employment patterns between two time points: when their child is in her first year of primary school (t_1) and the year before (t_0). Regarding mothers’ employment and work hours we also refer to outcomes measured in the latter period (t_0) as *pre-treatment* outcomes. We use information on mothers whose child entered primary school between 1999 and 2013. Mothers can appear more than once in our sample if they have several children who enter primary school during those years. Our estimations rely on 4,254 observations, which relate to 3,258 different mothers. All presented standard errors are clustered at the level of the mother and, hence, take multiple observations into account. Moreover, the robustness section shows that our results are robust to only using one observation per mother. The 1999-2013 time window is chosen because the 1999 school cohort is the first to have fully benefited from the 1997 legal right to a subsidised kindergarten place starting at the age of three (see Bauernschuster and Schlotter, 2015).⁷ There are also data reasons: some of our control variables regarding

⁷Although starting to observe children from 1999 onward precedes the expansion of afternoon care, we still observe children in both treatment and control groups between 1999 and 2003. Few children receive afternoon care between 1999 and 2003; this increases from 2003 onwards. Therefore, we also conduct a sensitivity analysis and look only at mothers and children observed from 2003 onward. The results remain very similar.

preschool are only regularly surveyed starting in 1999. Similarly, we can only include mothers whose child enters school in 2013 or before, because information on more recent cohorts is not yet available. In the robustness section, we experiment with a shorter time period.

In order to get closer to an ideal experimental situation, we only consider respondents who were interviewed between January and July in both years (t_0 and t_1). This ensures that the employment patterns observed at t_0 and t_1 , respectively, precedes and follows enrolment in an afternoon care service, which takes place at the start of the school year in August/September. As the vast majority of interviews in SOEP takes place in the first half of the year, these two sample restrictions reduce the sample size by only 6.2%.⁸ We only examine school starters for two reasons. First, there is a high degree of persistence in the participation in all-day schooling (Steiner, 2011). Hence, there is little variation in a child’s treatment status over time. Second, as the school entry date is rather exogenous to the individual family, we argue that reverse causality issues are much more a concern for children who change their treatment status while they are already in school.

We omit individuals with missing information on the key variables (maternal labour force status in t_0 and/or t_1 , child’s afternoon care status in t_1). This reduces the sample size by 4%. However, we include individuals with missing values in control variables by using a separate missing-value dummy for each variable with missing values. In a robustness test, we show that disregarding these observations does not change our conclusions. As a final sample restriction, we only look at children who turn 5, 6, or 7 in the year of their school entry. This restriction reduces the sample by 1.2% and is imposed in order to reduce measurement error in school entry. Again, the results are robust to including these cases as well.

4.2. *Treatment and control group*

Our control group includes mothers with children who are only in school in the morning and who do not receive formal afternoon care. The treatment group includes mothers whose child is (i) in an all-day school until mid or late afternoon; or (ii) attends primary school and a separate after-school programme (*Hort*). As mentioned earlier, the two forms of provision cannot be easily distinguished, as all-day schools typically offer afternoon activities run by staff external to the school, thus closely resembling afternoon programmes provided by social services and non-profit organisations, either on school premises or in

⁸The results are also robust to using cases with an interview between August and December; see section 6.

other care centres. In addition, before 2009 the SOEP questionnaire does not allow for differentiating between the two types of afternoon programmes.⁹ While this is problematic for studies trying to separate all-day schools and after-school programmes, it does not affect our strategy. The precise administrative structure of the afternoon care is not relevant to maternal employment; what matters is whether the child attends a formal afternoon service, which is what we capture. In 2009 the relevant SOEP questions changed to unambiguously classify students according to the type of afternoon care in which they participate. We use this information and, as a robustness check, we exclude children attending *Hort* from the analysis.

In summary, our sample consists of mothers whose child attends day care in t_0 and primary school in t_1 . Mothers are assigned to the treatment group if in t_1 their child usually receives afternoon care, either defined as attendance at primary school the whole day or attending primary school and a separate after-school programme; the remaining mothers in the sample constitute the control group, whose children only attend primary school in the morning. Our sample consists of 4,254 mother-child pairs: 1,278 in the treatment group and 2,976 in the control group.

4.3. Outcome variables

Two indicators describe our maternal employment outcomes, one relates to the extensive margin and one to the intensive margin. The first outcome variable is binary and indicates whether a mother works in t_1 , while the second outcome indicates the number of actual hours a mother works per week in t_1 . This second outcome takes on the value 0 for individuals who are not working.

4.4. Conditioning variables

We make use of a broad range of control variables in our analyses. All originate from the interview in t_0 and, hence, describe the situation before the child enters primary school. We include control variables that are likely to be related to both maternal labour force participation in t_1 and child's attendance in afternoon services (see also the discussion in Section 3). These conditioning variables describe the labour market history of the mother, their education, and their demographic characteristics. For their labour market history, we do not just rely on information about their work status or weekly hours of work in

⁹Before 2009 the question reads "Which of the following institutions do [your] children currently attend?" and lists both primary school and after-school programme as category leaving parents the possibility to give multiple answers.

t_0 , but also on the same variables from two years before the child’s school entry and retrospective information on the number of years in full-time and part-time employment. In addition, the SOEP, unlike many other general surveys, elicits rich information on work preferences and plans. For mothers who are in employment at t_0 , we include the number of desired working hours, while, for mothers who are not in work at t_0 , we include variables on job search behaviour and working intentions.¹⁰

In addition to information about mothers, we also use information regarding their child, their partner, and the household. Child characteristics included are age, the presence of younger siblings, and enrolment in childcare in t_0 . The characteristics of maternal partners include labour market attachment and education as well as some demographics. Household information relates to income, rural/urban classification of the place of residence, federal state, as well as state level unemployment and GDP. Additionally, we include indicators for each survey year as well as for the subsamples of SOEP.

Table 1 shows the means of selected conditioning variables in t_0 for mothers in the treatment and control groups, respectively (Table A.1 in the appendix provides a full list of control variables). This table compares the means of the treatment group (column 1) with the means of the unmatched control group (column 2). The similarity between treatment and control groups is shown by the mean differences (column 3) - the difference in the mean between treatment and control groups. In addition, Table 1 comprises information of the means of the matched control group, i.e. re-weighted by entropy balancing (column 4), and the similarity between treatment and control groups is shown by the standardized bias (columns 5 and 6).

[Table 1 about here]

The descriptive comparison between treatment and control groups shown in Table 1 suggests that maternal labour supply differs already in t_0 between treatment and control group. Mothers whose children receive afternoon care are 18 percentage points more likely to have worked in t_0 and work about 11 hours more a week in t_0 , i.e. prior to their child’s school entry. Table 1 also points to substantial differences with respect to other

¹⁰We include binary variables for each answer (see Table A.1 in the appendix) to the questions “Do you intend to engage in paid employment (again) in the future?”, “When, approximately, would you like to start with paid employment?”, “Are you interested in full-time or part-time employment, or would both suit you?”, “Is it or would it be easy, difficult or almost impossible to find an appropriate position?”, “Could you start working within the next two weeks?”, and “Have you actively looked for work within the last four weeks?”.

characteristics between mothers whose child receives afternoon care and those whose child does not. For instance, mothers with children in afternoon care services are much more likely to live without a partner, to have a highest secondary school degree, and to have more full-time work experience. Moreover, their children are more likely to have attended day care for longer hours.

With regards to the standardized difference for the unmatched control group (see column 5 in Table 1), for many variables it exceeds the value of 20 in absolute terms, which is considered to be a large difference (Rosenbaum and Rubin, 1985). But Table 1 also shows that after re-weighting with entropy balancing weights, the matched control group has the same mean as the treatment group in all variables (see column 4) and a standardized bias of zero (column 6).¹¹

5. Results

Table 2 provides an initial snapshot of the differential changes in employment patterns between mothers whose child is enrolled in afternoon services and those whose child is in school only in the morning. Such differential change is most visible when mothers are grouped according to their employment pattern the year before their child enters school (t_0). We start by looking at mothers who are not in paid work in t_0 . We then divide this group according to their children’s afternoon services attendance in t_1 and notice that only 65% of those whose child is cared for in the afternoon are not in work, as opposed to 79% of those whose child is not. While both groups of mothers are more likely to work part-time than full-time, 7% of those in the treatment group work full-time, while only 1.6% do so among those whose child does not attend afternoon care.

[Table 2 about here]

When looking at mothers who are already in paid work the year before school starts, differences between those using afternoon care and those who do not are less stark and relate to working hours rather than employment status. For example, among those who work part-time before their child enters primary school, 8% shift to full-time work if their

¹¹In Table A.1 we show that propensity score matching also works well in reducing the differences between treatment and control groups. None of the standardized biases are larger than 20 % after propensity score matching; although several values are greater than 5%, which is considered to be a threshold for low values (see Caliendo and Kopeinig, 2008). However, the standardized bias is clearly smaller for the entropy balancing specification than for the propensity score specification. For some variables, such as child’s gender, the standardized bias in the propensity score specification is even larger than in the unweighted control group.

child receives afternoon care, as opposed to only 3.8% in the control group. It is also notable that among mothers who are working full-time before their child is in school, 17% scale back to part-time if their child is in half-day school, as opposed to only 10% of those whose child attends an afternoon care service. Overall, mothers whose child attends school for a full day or combined with an after-school programme are more likely to start working and to work longer hours than mothers whose child is only attending school for half a day. This difference is visible for all groups of mothers irrespective of the employment status before the child’s entry in school. In the following we will continue to differentiate the results according to mothers’ t_0 working status.

The regression results in the first column of Table 3 confirm these conditional correlations for all mothers (Panel A), as well as for mothers who did not work in t_0 (Panel B) and those who worked in t_0 (Panel C). This column is based on our baseline regression, which controls only for the value of the outcome variable when the child was in kindergarten (i.e. the t_0 working status and actual working hours, respectively), alongside time and state fixed effects (capturing general differences over time and between states). These conditional correlations do not imply causality, as mothers in treatment and control group have different socio-economic and demographic characteristics that might underlie the differential changes in the outcome. For instance, mothers in the treatment group are, on average, better educated (see Table 1) and this labour market advantage might make them more likely to start working or to increase their working hours (even conditional on their previous labour market status). Hence, the conditional correlations presented would be (upward-) biased estimates of the true causal effects.

We take into account differences in the observed characteristics in the second column of Table 3. This specification is based on the regression-adjusted matching procedure with entropy balancing outlined in Section 3. Column (2) suggests that the child’s participation in afternoon care increases maternal employment and maternal working hours, irrespective of the mother’s t_0 working status. The coefficients in this specification are of similar magnitude as in the regression without control variables in column (1). Column (3) further exploits the richness of our data by including a set of variables on working preferences and intentions, elicited at t_0 . This way we are comparing mothers who not only have similar characteristics and behave similarly, but who also express similar preferences in relation to work. For all mothers, we find an increase in employment of 7.5 percentage points, compared to the mean of 72.5% of mothers who worked prior to school entry of their child, this effect of afternoon care participation increases the share of working mothers by 10 percent. For mothers who did not work prior child’s entry to school, their child’s

afternoon care increases the probability of taking up paid work by 11.4 percentage points. Our effect size for mothers who did not work prior to school entry of their child is very similar to what Berthelon et al. (2015) find for Chilean mothers: The increase in full-day schooling in Chile increases mothers' labour supply by 11.9 percentage points.

Among mothers who are already in work during the preschool years, the effect of the child being in afternoon care is less pronounced on the mother's decision to work or not, but still significant (+5.4 percentage points). Further, these mothers take advantage of afternoon care to increase their weekly working time by 2.6 hours on average. With respect to the intensive margin, our findings are related to Felfe et al. (2016) who find that Swiss mothers whose children participate in afternoon care are more likely to work full-time (3.3 percentage points). However, our results differ to those of Felfe et al. (2016), as they do not find an effect on take up of employment in their study. This might be due to their use of different age groups of children. Our results are comparable to studies exploiting the expansion of all-day schools in Germany. Both Nemitz (2015) and Shure (2016) find positive effects on maternal employment but no effect on the number of hours worked. In particular, the results by Shure (2016) are in line with ours, as the study reports an increase of five percentage points in mothers' probability to work. Note that she uses school level data rather than county level ones, but in doing so is only able to include data from four German states. By contrast Dehos and Paul (2017) find no effects of afternoon care participation on maternal labour supply both at the intensive and the extensive margin. However, as discussed before, it is unclear whether the expansion of afternoon care programmes as measured at the county level can indeed be seen as exogenous. With regards to other effects on labour supply, our estimates are comparable to studies analysing the impact of family policies in Germany on maternal labour supply: Geyer et al. (2015) show that the provision of childcare for children from age one onwards together with the parental leave reform in 2007 increases employment for mothers with very young children by 7 percentage points.

[Table 3 about here]

The main message from these results is that a child's participation in afternoon care services seems to affect maternal labour supply, both on the extensive and intensive margin. Women who are not employed during their child's preschool years are more likely to take up paid work if their child attends afternoon care. On the other hand, mothers who are already in employment before their child starts school appear able to extend their working hours and remain employed.

However, it remains a question of whether afternoon care is replacing other forms of non-formal childcare, thus truly affecting maternal employment but crowding out unpaid childcare by relatives and friends or paid childcare by childminders. We explore this hypothesis using information about other types of childcare used. We use the same specifications as for maternal employment shown in Table 3. Only this time our outcomes are a series of binary indicators measuring whether other forms of non-formal childcare are used in the first year of school (t_1). We find very little evidence of any substitution between afternoon care services and other forms of non-formal childcare (Table 4). Interestingly, Table 4 shows an increase in the reliance on friends among mothers following the child’s participation in afternoon care services. This might reflect that opening hours of afternoon care services do not completely account for a full working day.

[Table 4 about here]

Next, we examine the effect of afternoon care on paternal employment. As fathers have increased their involvement in childcare over time, it is possible that the effects of extended school days ripple into their working pattern. We find no evidence of this, thus confirming the in-elasticity of paternal labour supply to childcare (Table 5).

[Table 5 about here]

As a last step, we examine whether the estimated average effects hide relevant differences between subgroups. Table 6 shows results across different subgroups, defined by either socio-economic status or family composition. All estimated coefficients reported are positive and most are statistically significant, suggesting that the detected increase in maternal labour supply is consequential to the child’s participation in afternoon care services and is not driven by one group in particular.¹² Within this general pattern, there are some differences between subgroups worth noting, although none of them are statistically significant.¹³ There is some evidence that more educated mothers are more likely to start working because of their child’s afternoon care participation, in line with the expectation that mothers with better employment prospects benefit more from afternoon care services than mothers with relatively less education. The pattern holds irrespective

¹²For the analyses in Table 6, we split our sample according to different characteristics of mother and child and run separate regressions in each subsample. This analysis includes control variables only in the regression step and not in the matching step as cell sizes in some subgroups become too small for the matching procedure.

¹³This is partly due to smaller sample size.

of the indicator used to define the “more educated” group, whether at least 12 years of education (columns 1 and 2) or holding a post-secondary degree (columns 3 and 4). By contrast, there are no clear differences in the effect sizes between mothers with and without migration background (columns 5 and 6). In relation to family composition, we find some evidence for differential effects according to the presence of younger children. Effects are less strong in Panel A if the child entering school has any younger sibling (columns 7 and 8) and, in particular, a sibling aged three or younger (columns 9 and 10). Further, there is some evidence in Panel A for stronger effects for mothers whose child had not attended an early education centre for the full day (columns 11 and 12). There is, however, little evidence for differences between cohorts of younger and older mothers (columns 13 and 14). Results on hours worked, reported in Panel B, are rather similar across the different groups, suggesting no heterogeneous effect. In sum, there is little evidence that our treatment effect is driven by a specific group of mothers. Instead, entering into paid employment or increasing working hours thanks to afternoon care appears to be a general behaviour common to different demographic groups with some evidence that more educated mothers may take greater advantage of afternoon care.

[Table 6 about here]

6. Robustness checks

In this section, we provide additional evidence for the robustness of our main results. We start by investigating the two main sources of bias that could potentially undermine our identification strategy: reverse causality (section 6.1) and omitted variable bias (section 6.2). At the end of this section, we also test the robustness of our results to applying different sample restrictions and estimation techniques (section 6.3).

6.1. Reverse causality

Reverse causality is an issue if the child is receiving afternoon care *because* the mother increases her labour force participation. We focus on school starters in our main specification in order to mitigate concerns of reverse causality as the school entry date is rather exogenous to the individual family. To further address this threat to our identification strategy, this section makes use of a specific feature of all-day schools: Parents must decide whether or not their child participates in the afternoon programme at the start of a school year.¹⁴ Hence, reverse causality is less of an issue in the case of

¹⁴This rule applies to “open all-day schools” only. These are those where participation to afternoon activities is voluntary and they constitute the large majority of all-day schools among primary schools

mothers who find a job after the beginning of the school year because they can only enrol their child to all-day schooling in the next school year.

Columns (2) and (3) of Table 7 present the results relating to these reverse causality considerations. In a first step, we demonstrate that the estimated effect sizes are rather similar if the treatment group consists exclusively of children who attend all-day schools (model 2). This model omits children who attend an after-school programme, *Hort*, as the aforementioned institutional peculiarity mainly relates to all-day schools. In a second step, we exclude all mothers from the sample of model (2) who started, quit, or changed their job after the interview in t_0 and before the school start in September of that year. This way we drop those mothers for whom the temporal ordering of events is clearly employment change first and all day school enrolment after.¹⁵ The estimated effects in model (3) are very similar to the effects in our main specification and in model (2), suggesting that this kind of reverse causality does not drive our results.

While these specifications consider the case of reverse causality resulting from mothers who change their labour supply before the child enters school, reverse causality might also arise from mothers who *plan* to change their labour supply after their child's school entry. This second kind of reverse causality is not directly taken into account in Table 7. However, note that the conditioning variables comprise several variables that strongly relate to planned changes in labour supply, like working intentions and job search behaviours for non-working individuals and desired working hours for working individuals. As long as planned changes in labour supply are picked up by these conditioning variables, our empirical strategy also addresses this second kind of reverse causality.

[Table 7 about here]

6.2. Omitted variable bias

To identify the effect of afternoon care on maternal labour supply, the models must include all variables affecting both afternoon care attendance and changes in maternal labour market participation. While, so far, all the control variables originate from t_0 , specific events might occur between t_0 and t_1 that might affect both mothers' working patterns and changes in children's participation in afternoon care services and, therefore,

(Marcus et al., 2013). In the 5% of schools, where participation in all-day schooling is compulsory for all the pupils, parents do not have the choice option at the beginning of the school year.

¹⁵This specification is not our preferred one as the restriction might be overly conservative: mothers know their child's treatment status before school start and, hence, can adapt their employment pattern in anticipation of the treatment.

constitute omitted variables. Hence, in model (4) of Table 7 we try to control for such events, which include the birth of a child between t_0 and t_1 , the absence of a partner in t_1 , and the partner’s involuntary job loss between t_0 and t_1 . We also consider the existence of various alternative childcare arrangements in t_1 . We do not include the t_1 variables in our main specification as they could also be affected by the treatment, and, hence, be bad control variables. It is reassuring to see that the inclusion of these additional controls does not alter the results. Further, in model (5) we additionally include state-year fixed effects in the regression step in order to take into account differential rates of expansion of afternoon care programmes across German states. Controlling for state-year effects does not change our findings, indicating that there is variation in afternoon care availability also within a given state in a given year.

Nevertheless, we cannot be sure that we included all relevant control variables. To assess how big the influence of potentially omitted variables must be in order to completely explain the obtained effects of afternoon care, we apply the method proposed by Oster (2013, 2016). This method builds on the idea that the bias from observed variables is informative regarding the bias from omitted variables (under the assumption that there is some kind of proportionality between the forms of bias). Oster (2016) elaborates on the approach suggested by Altonji et al. (2005) and the often applied procedure of looking at coefficient stability after the inclusion of control variables. The main contribution of her method is to take into account the explanatory power (and thereby the relevance) of the included control variables: The method relates the change in the estimated treatment effect (due to the inclusion of control variables) to the associated change in the R^2 . More formally, Oster (2016) approximates a bias-adjusted treatment effect, β^* , by

$$\beta^* \approx \tilde{\beta} - \tilde{\delta} [\dot{\beta} - \tilde{\beta}] \frac{(R_{max}^2 - \tilde{R}^2)}{(\tilde{R}^2 - \dot{R}^2)}, \quad (2)$$

where $\dot{\beta}$ and \dot{R}^2 are the estimated treatment effect and coefficient of determination from a baseline regression without additional control variables and $\tilde{\beta}$ and \tilde{R}^2 are their equivalents from the full regression with additional control variables. While all these four quantities can be estimated from the data, we need to make some assumptions regarding $\tilde{\delta}$ and R_{max}^2 . $\tilde{\delta}$ is assumed to be positive and denotes the degree of proportionality. It indicates how much of the variation in the outcome is explained by the observed controls versus unobserved. $\tilde{\delta} = 1$ means that we assume an equal importance of observed and unobserved factors (“equal selection assumption”), while $\tilde{\delta} > 1$ [$\tilde{\delta} < 1$] indicates that the degree of selection on unobserved variables necessary to explain away the effects is stronger [weaker]

than selection on the observables. R_{max}^2 denotes the share of variation in the outcome variable that is explained by observed and unobserved variables together. It is less than one if there is measurement error in the dependent variable. Oster (2013) derived a value for R_{max}^2 based on empirical reasoning and suggests $R_{max}^2 = \min\{2.2 \cdot \tilde{R}^2, 1\}$.¹⁶

Based on Eq. (2), Oster (2016) suggests two closely related approaches to evaluate the robustness to omitted variable bias. We follow the suggestions and report in Table 8 (i) a lower bound of the treatment effect assuming equal selection on observed and unobserved variables (i.e. we set $\tilde{\delta} = 1$); and (ii) the degree of proportionality for which our treatment effect would equal zero (i.e. we set $\beta^* = 0$). While the first approach checks whether the lower bound is still larger than 0 and included in the 95% confidence interval of the previously estimated treatment effect, the second approach examines whether $\tilde{\delta} > 1$, i.e. if selection on unobserved variables has to be more important than selection on observables in order to pull the estimated effect of afternoon care to zero. This would be the case if the unobserved variables are more important than the whole set of included control variables, which we selected drawing on the literature on maternal employment and childcare.

The first two columns in Table 8 basically repeat the estimated treatment effects from the baseline and the main specification (see also Table 3), while the other columns display the results of the two approaches outlined above. Panel A shows that in the baseline model, afternoon care increases the probability of working by 15.4 percentage points for mothers who are not working in t_0 . The inclusion of the full set of control variables in our main model leads to a decrease in the treatment effect of about 4 percentage points but to an increase in R^2 from 0.04 to 0.38.¹⁷ Based on these estimates and Eq. (2), we calculate that $\tilde{\delta}$ would have to be as large as 1.91 in order to completely explain the estimated effect.¹⁸ This means that the influence of omitted variables needs to be almost twice as important as of the observed factors included in the model to bring the effect of afternoon care to zero. The estimated lower bound of the treatment effect is 0.058. It is larger than 0 and included in the 95% confidence interval around the estimated treatment effect. For mothers who worked in t_0 (Panel B), our estimates of $|\delta|$ are also both greater than 1 (and the values for $|\delta|$ even exceed that of Panel A). Further, the lower bounds are clearly larger than 0 and included in the respective confidence bands, suggesting that it

¹⁶Note that in the published article, Oster (2016) suggests a value of $R_{max}^2 = \min\{1.3 \cdot \tilde{R}^2, 1\}$. Nevertheless, we rely on the working paper version, which is more conservative as it generally generates higher values of R_{max}^2 .

¹⁷Similarly the Adjusted R^2 increases from 0.03 to 0.33.

¹⁸We use the Stata command *psacalc* provided by Oster (2013) to calculate the estimates of δ and the lower bound.

is very unlikely that our findings are explained by omitted variable bias.¹⁹ In this case, the value for the lower bound is calculated based on $\tilde{\delta} = -1$. Also in this case the lower bound is clearly within the 95% confidence band around $\tilde{\beta}$.

Thus, the method developed by Oster (2016) corroborates our findings.

[Table 8 about here]

6.3. *Alternative sample restrictions and estimation methods*

The next set of specifications addresses different sample restrictions (see Table 9). Model (1) disregards all observations with missing information on the control variables due to item non-response (see section 4.4). Specification (2) extends the sample by lifting the restrictions on the month of the interview and the age of the child (see section 4.1), while model (3) restricts the sample to a shorter observation period, the years 2003-2013. 2003 is the year in which a federal investment programme was launched to foster the expansion of all-day schools. In model (4) we only include one observation for each mother, namely the observation that refers to the mother’s first child to enter primary school in our observation period. Table 9 shows that our findings are robust to the different sample restrictions.

In the last set of sensitivity checks, we assess a number of issues regarding our estimation method. Model (5) presents the results obtained from Ordinary Least Squares estimation with the same set of control variables, while column (6) displays the estimates from propensity score matching.²⁰ Specification (7) performs entropy balancing separately according to the mother’s working status in t_0 . Finally, as there is an ongoing discussion about whether one should apply survey weights in matching applications or not (Solon et al., 2015), we re-estimate our main specification using survey weights in both the entropy balancing and the regression step (see column 8). Further, the results are very similar, when we use Tobit regression or Hurdle regression (Cragg, 1971) in the regression step (see Table A.2 in the Appendix). Our findings are robust to all these sensitivity checks.

[Table 9 about here]

¹⁹Note that the $\tilde{\delta}$ found for the outcome “working” in Panel B is negative as the treatment effect moves away from zero rather than toward zero when including control variables. The large and negative value of $\tilde{\delta}$ implies that any bias in the estimated effect due to omitted variables would have to be not only substantially larger than the bias generated by omitting observable variables, but it would also have to be in the opposite direction of the bias from omitting observable variables.

²⁰For the propensity score matching, we rely on kernel matching with a Gaussian kernel and a bandwidth of 0.06 (see Heckman et al., 1997; Marcus, 2014). We rely on the user-written programme “psmatch2” in Stata (Leuven and Sianesi, 2003).

7. Conclusion

In this paper, we examine how maternal labour supply changes due to her child's participation in afternoon care. A vast literature in economics provides evidence on the effect of early childhood care services provision on maternal employment, yet little is known about effects on maternal employment of subsidized childcare for school-aged children. One might argue that the problem of arranging for suitable childcare clearly becomes less pressing once children reach school age. Yet the problem does not disappear, as primary school children are not capable of self-care and full-time working hours rarely match the school day hours. Such mismatch is especially visible in countries with half-day schools, whereby children return home at lunch time. Germany is a case in point and the example we examine. We extend the existing literature by focusing on primary school children and examine how hours spent in formal afternoon care affect maternal labour supply patterns. We consider not only labour market participation, but also the actual hours worked, thus covering variations both at the extensive and intensive margin.

Our empirical strategy combines pre-treatment outcome modelling and matching, in which the treatment group consists of mothers whose children participate in formal afternoon care (i.e. in all-day schooling and/or after-school programmes). We use a non-parametric matching technique, entropy balancing, to generate a control group of mothers with similar characteristics whose children do not participate in afternoon care. The matching procedure considers a wide range of control variables, including the mother's detailed labour market history, the child's attendance in preschool, several characteristics of the partner, as well as household, regional labour market characteristics, as well as some often unobserved information like job search intentions and desired working hours. Our identification strategy is robust against selection on observables as well as against selection on unobserved variables with time-constant effects. We assume that conditional on the mother's labour force status before the child's school entry, there are no other variables other than the included control variables that simultaneously affect the child's participation in afternoon care and the mother's labour force status when the child is in first grade. We evaluate the robustness of our results to bias resulting from potentially omitted variables applying the method developed by Oster (2013, 2016).

Across the whole sample, we find that the child's being at school in the afternoon increases the mother's probability to start working, to remain working, and to increase the number of hours they work as their child enters school. Splitting the sample according to maternal work status prior to their child's school entry shows that the child's participation in afternoon care increases the likelihood of mothers who did not work before to take

up paid work by 11.4 percentage points. Furthermore, mothers who already worked during the year prior to their child’s school enrolment increase their working hours by an average of 2.6 hours per week due to their child’s participation in afternoon care services. In general, these results are in line with studies on childcare availability for children below compulsory school age in Germany (Bauernschuster and Schlotter, 2013) and other countries (e.g. Berlinski and Galiani, 2007) as well as with the few studies on school-aged children in other countries (Berthelon et al., 2015; Felfe et al., 2016). We do not find any effects on paternal labour force participation, underlining that maternal and paternal labour force participation differs, as the greater responsibility of mothers for children lead them to interrupt or reduce their labour force participation.

While our findings pertain to the specific case of Germany, nevertheless they draw attention to a more general issue: childcare availability continues to shape maternal employment patterns well after school entry. While so far the focus of researchers and policy-makers alike has mainly been on pre-school children, our analysis highlights that the need for childcare does not end when the child enters school. Policy-makers intending to foster maternal labour force participation should improve childcare opportunities not only for pre-school children but also for young school-aged ones.

Conflict of Interest Statement

The authors declare that they have no conflict of interest.

References

- Allemann-Ghionda, C. (2009). Ganztagsschule im europäischen Vergleich. Zeitpolitiken modernisieren - Durch Vergleich Standards setzen? *Zeitschrift für Pädagogik*, Beiheft 54:190–208.
- Altonji, J. G., Elder, T. E., and Taber, C. R. (2005). Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools. *Journal of Political Economy*, 113:151–184.
- Autorengruppe Bildungsberichterstattung (2016). Bildung in Deutschland 2016. Ein indikatorengestützter Bericht mit einer Analyse zu Bildung und Migration. W. Bertelsmann Verlag, Bielefeld.
- Baker, M., Gruber, J., and Milligan, K. (2008). Universal child care, maternal labour supply, and family well-being. *Journal of Political Economy*, 116(4):709–745.
- Bang, H. and Robins, J. (2005). Doubly robust estimation in missing data and causal inference models. *Biometrics*, 61(4):962–973.
- Bauernschuster, S. and Schlotter, M. (2013). Public child care and mothers’ labor supply - evidence from two quasi-experiments. *CESifo Working Paper Series 4191*.
- Bauernschuster, S. and Schlotter, M. (2015). Public child care and mothers’ labor supply - evidence from two quasi-experiments. *Journal of Public Economics*, 123(C):1–16.
- Beblo, M., Lauer, C., and Wrohlich, K. (2005). Ganztagsschulen und Erwerbsbeteiligung von Müttern: Eine Mikrosimulationsstudie für Deutschland. *Zeitschrift für ArbeitsmarktForschung - Journal for Labour Market Research*, 38(2):357–372.
- Berger, L. M., Hill, J., and Waldfogel, J. (2005). Maternity leave, early maternal employment and child health and development in the US. *The Economic Journal*, 115(501):F29–F47.
- Berlinski, S. and Galiani, S. (2007). The effect of a large expansion of pre-primary school facilities on preschool attendance and maternal employment. *Labour Economics*, 14(3):665–680.
- Berthelon, M., Kruger, D., and Oyarzun, M. (2015). The effects of longer school days on mothers’ labor force participation. *IZA Discussion Paper*, 9212.
- Bettio, F. and Plantenga, J. (2004). Comparing care regimes in Europe. *Feminist Economics*, 10(1):85–113.

- Blau, D. M. and Currie, J. (2006). Preschool, day care, and after-school care: Who’s minding the kids? In Welch, F. and Hanushek, E. A., editors, *The Handbook of Economics of Education*, pages 1163–1267. North-Holland, New York.
- Brilli, Y., Del Boca, D., and Pronzato, C. (2013). Does child care availability play a role in maternal employment and children’s development? Evidence from Italy. *Review of Economics of the Household*, pages 1–25.
- Caliendo, M. and Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys*, 22(1):31–72.
- Cascio, E. (2009). Maternal labor supply and the introduction of kindergartens into American public schools. *Journal of Human Resources*, 44(1):140–170.
- Cragg, J. G. (1971). Some statistical models for limited dependent variables with application to the demand for durable goods. *Econometrica*, 39(5):829–844.
- Daly, M. (2000). *The gender division of welfare: The impact of the British and German welfare states*. Cambridge University Press, Cambridge.
- Dehos, F. and Paul, M. (2017). The effects of after-school programs on maternal employment. *SOEPpapers on Multidisciplinary Panel Data Research*, (905).
- Federal Ministry of Education and Research (2016). Was ist der Unterschied zwischen offenen und gebundenen Ganztagschulen? http://www.ganztagschulen.org/archiv/188_306.php. Accessed on 10/07/16.
- Felfe, C., Lechner, M., and Thiemann, P. (2016). After-school care and parents’ labour supply. *Labour Economics*, 42(3):64–75.
- Fitzpatrick, M. (2012). Revising our thinking about the relationship between maternal labor supply and preschool. *Journal of Human Resources*, 47(3):583–612.
- Gelbach, J. (2002). Public schooling for young children and maternal labor supply. *American Economic Review*, 92(1):307–322.
- Geyer, J., Haan, P., and Wrohlich, K. (2015). The effects of family policy on maternal labor supply: Combining evidence from a structural model and a quasi-experimental approach. *Labour Economics*, 36:84–98.
- Gornick, J. C. and Meyers, M. K. (2003). *Families that work: Policies for reconciling parenthood and employment*. Russell Sage Foundation.

- Goux, D. and Maurin, E. (2010). Public school availability for two-year olds and mothers' labour supply. *Labour Economics*, 17(6):951–962.
- Hagemann, K. (2006). Between ideology and economy: The time politics of child care and public education in the two Germanys. *Social Politics: International Studies in Gender, State & Society*, 13(2):217–260.
- Hainmueller, J. (2012). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis*, 20(1):25–46.
- Hainmueller, J. and Xu, Y. (2013). Ebalance: A stata package for entropy balancing. *Journal of Statistical Software*, 54(7):1–18.
- Havnes, T. and Mogstad, M. (2011). No child left behind: Subsidized child care and children's long-run outcomes. *American Economic Journal: Economic Policy*, 3(May 2011):97–129.
- Heckman, J., Ichimura, H., and Todd, P. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *Review of Economic Studies*, 64(4):605–654.
- Holtappels, H. G., Klieme, E., Rauschenbach, T., and Stecher, L., editors (2008). *Ganztagsschule in Deutschland: Ergebnisse der Ausgangserhebung der "Studie zur Entwicklung von Ganztagsschulen (StEG)"*. Weinheim: Juventa Verlag.
- Joshi, H., Macran, S., and Dex, S. (1996). Employment after childbearing and women's subsequent labour force participation: Evidence from the British 1958 birth cohort. *Journal of Population Economics*, 9(3):325–348.
- KMK (2016). Allgemeinbildende Schulen in Ganztagsform in den Ländern in der Bundesrepublik Deutschland - Statistik 2010 bis 2014. *Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland*. <https://www.kmk.org/dokumentation-und-statistik/statistik/schulstatistik/allgemeinbildende-schulen-in-ganztagsform.html>.
- Knittel, T., Henkel, M., Krämer, L., Lopp, R., and Schein, C. (2014). *Dossier Müttererwerbstätigkeit: Erwerbstätigkeit, Erwerbsumfang und Erwerbsvolumen 2012*. Prognos AG, Berlin.
- Kosonen, T. (2014). To work or not to work? The effect of childcare subsidies on the labour supply of parents. *The BE Journal of Economic Analysis & Policy*, 14(3):817–848.
- Kullback, S. (1959). *Information theory and statistics*. New York: Dover Publications.

- Lange, J. (2015). "Da war doch noch was?" Der Hort als wenig beachtete Betreuungsalternative zur Ganztagsschule im Grundschulalter. *Kommentierte Daten der Kinder- und Jugendhilfe*, 18:9–11.
- Lange, J. (2016). Der Hort: viel genutzt, wenig beachtet! *DJI Impulse*, 2:21–23.
- Lefebvre, P. and Merrigan, P. (2008). Child-care policy and the labor supply of mothers with young children: A natural experiment from Canada. *Journal of Labor Economics*, 26(3):519–548.
- Leuven, E. and Sianesi, B. (2003). PSMATCH2: Stata module to perform full mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing. <http://ideas.repec.org/c/boc/bocode/s432001.html>. This version 4.0.11 22oct2014.
- Lewis, J., Campbell, M., and Huerta, C. (2008). Patterns of paid and unpaid work in Western Europe: Gender, commodification, preferences and the implications for policy. *Journal of European Social Policy*, 18(1):21–37.
- Lundin, D., Mörk, E., and Öckert, B. (2008). How far can reduced childcare prices push female labour supply? *Labour Economics*, 15(4):647–659.
- Marcus, J. (2014). Does job loss make you smoke and gain weight? *Economica*, 81(324):626–648.
- Marcus, J., Nemitz, J., and Spiess, C. K. (2013). Ausbau der Ganztagsschule: Kinder aus einkommensschwachen Haushalten im Westen nutzen Angebote verstärkt. *DIW Wochenbericht*, 27.
- Marcus, J., Nemitz, J., and Spieß, C. K. (2016). Veränderungen in der gruppenspezifischen Nutzung von ganztägigen Schulangeboten - Längsschnittanalysen für den Primarbereich. *Zeitschrift für Erziehungswissenschaft*, 19(2):415–442.
- Mattes, M. (2011). Children, families and states. Time policies of childcare, preschool and primary education in Europe. chapter Economy and politics: The time policy of the East German childcare and primary school system, pages 344–363. Oxford: Berghahn Books.
- Morgan, K. J. (2005). The "production" of child care: How labor markets shape social policy and vice versa. *Social Politics: International Studies in Gender, State & Society*, 12(2):243–263.
- Nemitz, J. (2015). The effect of all-day primary school programs on maternal labor supply. ECON – Working Papers 213, Department of Economics, University of Zurich.

- Nollenberger, N. and Rodríguez-Planas (2015). Full-time universal childcare in a context of low maternal employment: Quasi-experimental evidence from Spain. *Labour Economics*, 36:124–136.
- OECD (2011). *Doing Better for Families*. OECD, Paris.
- OECD (2015). *Education Policy Outlook 2015: Making Reforms Happen*. OECD, Paris.
- Oster, E. (2013). Unobservable selection and coefficient stability: Theory and validation. NBER Working Paper No. 19054, NBER, Cambridge Massachusetts.
- Oster, E. (2016). Unobservable selection and coefficient stability: Theory and validation. *Journal of Business Economics and Statistics*, forthcoming.
- Plantenga, J. and Remery, C. (2017). Out-of-school childcare: Exploring availability and quality in EU member states. *Journal of European Social Policy*, 27(1):25–39.
- Rosenbaum, P. and Rubin, D. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *American Statistician*, 39:33–38.
- Schober, P. S. and Spiess, C. K. (2015). Local day care quality and maternal employment: Evidence from East and West Germany. *Journal of Marriage and Family*, 77(3):712–729.
- Shure, N. (2016). School hours and maternal labour supply: A natural experiment from Germany. Department of Quantitative Social Science Working Paper 16-13, Institute of Education, London.
- Sigle-Rushton, W. and Waldfogel, J. (2007). Motherhood and women’s earnings in Anglo-American, Continental European, and Nordic countries. *Feminist Economics*, 13:55–91.
- Simonsen, M. (2010). Price of high-quality daycare and female employment. *The Scandinavian Journal of Economics*, 112(3):570–594.
- Solon, G., Haider, S., and Wooldridge, J. (2015). What are we weighting for? *Journal of Human Resources*, 50(2):301–316.
- Steiner, C. (2011). *Ganztagsschule: Entwicklung, Qualität, Wirkungen*, chapter Teilnahme am Ganztagsbetrieb - Zeitliche Entwicklung und mögliche Selektionseffekte, pages 57–75.
- Wagner, G. G., Frick, J. R., and Schupp, J. (2007). The German Socio-Economic Panel Study (SOEP) - Scope, evolution, and enhancements. *Schmollers Jahrbuch*, 127(1):139–169.
- Waldfogel, J. (1998). The family gap for young women in the United States and Britain: Can maternity leave make a difference? *Journal of Labor Economics*, 16(3):505–545.

- Wiezorek, C., Stark, S., and Dieminger, B. (2011). "Wissen Sie, die Infrastruktur ist einfach nicht so, dass ich aus dem Vollen schöpfen kann" – Ganzttagsschulentwicklung in ländlichen Räumen. *Zeitschrift für Erziehungswissenschaft*, 15:109–124.
- Wunder, C. and Heineck, G. (2013). Working time preferences, hours mismatch and well-being of couples: Are there spillovers? *Labour Economics*, 24:244–252.

Table 1: Summary of selected conditioning variables for treatment and control groups

	Mean			Mean		Standard. Bias (%)	
	Afternoon care	No afternoon care (unmatched)	Mean diff.	No afternoon care (matched)	unmatched	matched	
	(1)	(2)	(3)	(4)	(5)	(6)	
<i>Maternal labour supply in t_0</i>							
Working	0.72	0.54	0.18***	0.72	38.3	0.0	
Actual working hours	22.27	11.45	10.81***	22.27	70.3	0.0	
<i>Maternal characteristics t_0</i>							
Migration background	0.19	0.28	−0.09***	0.19	−20.2	0.0	
Age mother	36.23	36.13	0.10	36.23	1.9	0.2	
No spouse	0.21	0.10	0.11***	0.21	32.0	0.0	
<i>School degree</i>							
Basic school	0.10	0.22	−0.12***	0.10	−31.6	−0.0	
Intermediate school	0.40	0.39	0.01	0.40	2.9	0.0	
Technical college	0.06	0.06	−0.00	0.06	−0.9	0.0	
Highest secondary	0.33	0.20	0.13***	0.33	28.9	0.0	
Other school	0.07	0.09	−0.02**	0.07	−8.8	0.0	
School drop-out	0.02	0.02	−0.01	0.02	−4.0	0.0	
<i>Work experience</i>							
Years part time	3.10	2.92	0.18	3.10	5.4	0.0	
Years full time	6.29	5.72	0.57***	6.29	11.1	0.0	
Work experience, missing	0.00	0.00	0.00	0.00	4.3	0.0	
<i>Child characteristics t_0</i>							
Attendance of ECEC centre full day	0.23	0.12	0.11***	0.23	30.1	0.0	
Younger siblings	0.40	0.46	−0.05***	0.40	−11.0	0.0	
Older siblings	0.42	0.55	−0.13***	0.42	−26.7	0.0	
Only child	0.27	0.14	0.13***	0.27	32.0	0.0	
Female child	0.48	0.50	−0.02	0.48	−4.1	0.0	
<i>Type of non-formal childcare (CC)</i>							
CC none	0.61	0.61	−0.01	0.61	−1.1	0.0	
CC relatives	0.28	0.25	0.03*	0.28	6.3	0.0	
CC friends	0.07	0.04	0.03***	0.07	11.2	0.0	
CC paid carer	0.05	0.03	0.02***	0.05	9.9	0.0	
<i>Household characteristics t_0</i>							
Home owner	0.41	0.56	−0.15***	0.41	−30.1	0.0	
Household income (in 1000)	48.12	48.64	−0.52	48.12	−1.3	0.0	
Unemployment share	9.85	7.78	2.07***	9.85	56.5	0.1	
<i>N</i>	1, 278	2, 976	4, 254				

Notes: This table displays descriptive statistics for selected conditioning variables from the reduced set of controls for treatment and control groups. Table A.1 provides the information for the complete list of conditioning variables. The first column presents the means for mothers whose children attend afternoon care (treatment group), the second column for mothers whose children do not participate in afternoon care (control group), and the third column comprises the mean differences between the two groups. Column four shows the mean of matched mothers in the control group, while columns five and six depict the percentage standardized bias for unmatched and matched conditioning variables.

Source: SOEP v31, significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. CC=childcare, ECEC=early childhood education and care

Table 2: Transition matrix

	Not working in t_1		Part-time in t_1		Full-time in t_1		
	No		No		No		
	Afternoon	afternoon	Afternoon	afternoon	Afternoon	afternoon	N
	care	care	care	care	care	care	
Not working in t_0	64.77	78.96	27.84	19.43	7.39	1.62	1711
	<i>-14.18***</i>		<i>8.41***</i>		<i>5.77***</i>		
Part-time in t_0	8.64	11.75	83.22	84.46	8.14	3.80	1998
	<i>-3.11*</i>		<i>-1.23</i>		<i>4.34***</i>		
Full-time in t_0	8.28	9.33	9.82	16.89	81.90	73.78	551
	<i>-1.05</i>		<i>-7.07*</i>		<i>8.12*</i>		

Notes: This table presents a transition matrix for the employment status of mothers when their child enters primary school, differentiated by treatment status. The numbers in the upper left cell indicate that 64.77 % of mothers who did not work in t_0 continue to not work in t_1 if their child is in afternoon care. The numbers shown in italics represent the percentage point differences between mothers whose child attends afternoon care and those whose children do not participate.

Source: SOEP v31, significance levels (based on robust standard errors clustered at the mothers' level): * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: The effect of afternoon care on maternal labour supply

	Baseline model	Reduced set of t_0 controls	Main model
Panel A: All mothers			
Working	0.081*** (0.014)	0.078*** (0.024)	0.075*** (0.024)
Hours	3.212*** (0.468)	2.808*** (0.858)	2.779*** (0.836)
N	4,254	4,254	4,254
Panel B: Not working in t_0			
Working	0.154*** (0.032)	0.128*** (0.039)	0.114*** (0.037)
N	1,711	1,711	1,711
Panel C: Working in t_0			
Working	0.053*** (0.015)	0.054** (0.025)	0.054** (0.025)
Hours	2.636*** (0.561)	2.554*** (0.943)	2.590*** (0.940)
N	2,543	2,543	2,543

Notes: Each cell depicts the effect of afternoon care participation on maternal labour supply indicators for different groups of mothers as indicated by the panel name. All regressions include state and time fixed effects. The first column comprises the association between after school care attendance and maternal labour supply controlling for the pre-treatment value of the respective outcome variable (in period t_0). The second column includes a reduced set of conditioning variables from t_0 , while the third column comprises the full set of conditioning variables. Both sets of variables are summarized in Table A.1 in the Appendix. The second and third column consider the weights from entropy balancing, while the first column is without matching.

Source: SOEP v31. Robust standard errors clustered at the mothers' level in parentheses, significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: The effect of afternoon care on other types of childcare

	Baseline model	Reduced set of t_0 controls	Main model
No child care	-0.027 (0.017)	0.005 (0.028)	0.003 (0.027)
Child care by relative	0.011 (0.015)	-0.009 (0.025)	-0.007 (0.025)
Child care by friend	0.032*** (0.009)	0.048*** (0.010)	0.050*** (0.009)
Paid child care	0.005 (0.010)	-0.020 (0.017)	-0.019 (0.017)
Child care, missing answer	0.004 (0.005)	0.002 (0.006)	0.002 (0.006)
<i>N</i>	4,254	4,254	4,254

Notes: Each cell depicts the effect of a child participating in afternoon care on binary indicators of other types of childcare, as indicated by the row name (see Table 3 for a description of the models).

Source: SOEP v31. Robust standard errors clustered at the mothers' level in parentheses, significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: The effect of afternoon care on paternal labour supply

	Baseline model	Reduced set of t_0 controls	Main model
Working	-0.006 (0.008)	0.016 (0.015)	0.015 (0.012)
Hours	-0.700 (0.509)	-0.044 (1.106)	-0.460 (0.936)
<i>N</i>	2,919	2,919	2,919

Notes: Each cell depicts the effect of a child participating in afternoon care on paternal labour supply (see Table 3 for a description of the models).

Source: SOEP v31. Robust standard errors clustered at the mothers' level in parentheses, significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Effect heterogeneity

Social status														
Education ≥ 12			Post-secondary degree		Migration background		Sibling composition							
no (1)	yes (2)		no (3)	yes (4)	no (5)	yes (6)	no (7)	yes (8)	no (9)	yes (10)	no (11)	yes (12)	no (13)	yes (14)
Panel A: Not working in t=0														
Working	0.090** (0.040)	0.173*** (0.055)	0.074 (0.059)	0.150*** (0.039)	0.098** (0.040)	0.127** (0.057)	0.184*** (0.061)	0.100*** (0.038)	0.155*** (0.048)	0.094** (0.043)	0.148*** (0.040)	0.067 (0.057)	0.129** (0.065)	0.113*** (0.038)
N	1035	664	454	1244	1175	536	686	1025	844	867	1308	384	790	921
p-value	0.186			0.234	0.645		0.207		0.315		0.181		0.819	
Panel B: Working in t=0														
Working	0.027 (0.030)	0.074*** (0.022)	0.025 (0.083)	0.047*** (0.017)	0.048** (0.019)	0.034 (0.046)	0.043** (0.021)	0.046 (0.032)	0.047** (0.020)	0.034 (0.038)	0.042* (0.023)	0.051* (0.028)	0.041** (0.021)	0.060*** (0.028)
p-value	0.183			0.751	0.761		0.940		0.731		0.806		0.569	
Hours	1.962* (1.022)	2.737*** (0.850)	4.341 (2.637)	1.988*** (0.644)	2.699*** (0.709)	1.414 (1.436)	2.408*** (0.770)	2.618** (1.087)	2.441*** (0.718)	3.077*** (1.360)	2.591*** (0.770)	2.390** (1.042)	2.104*** (0.797)	3.119*** (0.949)
N	1173	1357	284	2243	2017	526	1685	858	1927	616	1513	1008	1285	1258
p-value	0.542			0.286	0.378		0.867		0.654		0.870		0.392	

Notes: Each cell depicts the effect of afternoon care participation on maternal labour supply indicators from a separate OLS regression. The estimated coefficients originate from unweighted regressions, in which the sample is split in two based on the column header. The displayed p -values relate to two-sided tests for the equality of treatment effects in the respective subgroups.

Source: SOEP v31. Robust standard errors clustered at the mothers' level in parentheses, significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: Sensitivity checks I: Identification issues

Identification issues					
	Main model (1)	Only children in all-day schools (2)	No job change <i>prior</i> September (3)	Including information from period t_1 (4)	+ state- year FE (5)
Panel A: Not working in t_0					
Working	0.114*** (0.037)	0.088** (0.035)	0.107*** (0.036)	0.108*** (0.034)	0.116*** (0.039)
<i>N</i>	1,711	1,513	1,346	1,346	1,346
Panel B: Working in t_0					
Working	0.054** (0.025)	0.067** (0.032)	0.057* (0.030)	0.063** (0.028)	0.073** (0.032)
Hours	2.590*** (0.940)	2.861*** (1.069)	2.687*** (0.964)	3.078*** (0.873)	4.176*** (1.017)
<i>N</i>	2,543	1,938	1,815	1,815	1,815

Notes: Each cell depicts the effect of afternoon care participation on maternal labour supply indicators. All models are based on the main specification (repeated in the first column). The second column shows the effect of afternoon care only for children participating in all-day primary schools, i.e. excluding children attending after-school programmes (*Hort*) from the analysis. The third column further restricts the sample to mothers without job change between interview and the beginning of the school year, while column four additionally controls for selected t_1 variables in matching and regression step. Column five additionally includes state-year fixed effects in the regression step.

Source: SOEP v31. Robust standard errors clustered at the mothers' level in parentheses, significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 8: Assessing the potential bias of omitted variables

			Bounds of β		Proportionality	
	Baseline model	Main model	Lower bound	In 95%-c.i. band	δ	$ \delta > 1$
Panel A: Not working in t_0						
Working	0.154*** (0.032)	0.114*** (0.037)	0.058	✓	1.914	✓
R ²	0.04	0.38				
Adj. R ²	0.03	0.33				
Panel B: Working in t_0						
Working	0.053*** (0.015)	0.054** (0.025)	0.051	✓	-34.073	✓
R ²	0.02	0.25				
Adj. R ²	0.01	0.22				
Hours	2.636*** (0.561)	2.590*** (0.940)	1.967	✓	3.950	✓
R ²	0.48	0.52				
Adj. R ²	0.48	0.50				

Notes: The first and second column comprise the baseline and main effect of afternoon care participation on maternal labour supply, respectively. All regressions include the pre-treatment values of the outcome variable as well as state and time fixed effects. The second column additionally considers the full set of conditioning variables. Based on the approach outlined in Oster (2016), the third column shows the lower bound of β and the fourth column checks whether this value is within the 95% confidence interval of the treatment effect. The fifth column reports the value of proportionality δ and shows how strong the influence of unobserved factors has to be compared to the observed to pull the treatment effect to zero (main effect). The last column checks whether $|\delta| > 1$.

Source: SOEP v31. Robust standard errors clustered at the mothers' level in parentheses, significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 9: Sensitivity checks II: Sample restrictions and estimation issues

	Sample restriction				Estimation issues			
	w/o missing information (1)	Full sample (2)	2003 – 2013 (3)	One child (4)	Ordinary Least Squares (5)	Propensity Score Matching (6)	Separate (7)	Survey weights (8)
Panel A: Not working in t_0								
Working	0.105*** (0.039)	0.127*** (0.033)	0.145*** (0.039)	0.116** (0.050)	0.126*** (0.032)	0.117*** (0.033)	0.162*** (0.035)	0.163*** (0.042)
<i>N</i>	1,573	1,956	1,278	1,330	1,711	1,711	1,711	1,711
Panel B: Working in t_0								
Working	0.053** (0.025)	0.081*** (0.026)	0.048* (0.025)	0.055** (0.026)	0.045*** (0.017)	0.059** (0.024)	0.063** (0.029)	0.062** (0.029)
Hours	2.370** (0.946)	3.837*** (0.976)	2.164** (0.956)	3.186*** (0.870)	2.423*** (0.620)	2.721*** (0.919)	3.179*** (1.120)	2.451** (1.027)
<i>N</i>	2,417	2,850	2,087	1,931	2,543	2,543	2,543	2,543

Notes: Each cell depicts the effect of afternoon care participation on maternal labour supply indicators. As before, all models are based on the main specification. The first column shows the effect of afternoon care estimated only for those mothers who have non-missing information on all variables. The second column relaxes the sample restriction regarding interview dates, while the third column only considers mothers whose children enter school from 2003 onwards. The fourth column includes only the first observed child for each mother. The fifth column comprises estimates obtained from ordinary least squares, and the sixth column from propensity score matching. Column seven performs entropy balancing separately according to the t_0 working status of mothers, while column eight includes sample weights in both entropy balancing and regression step. Source: SOEP v31. Robust standard errors clustered at the mothers' level in parentheses, significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix

Table A.1: Descriptive statistics of reduced and full set of t_0 controls - before and after matching

Variable	Means treated	Means controls			Standard. Bias (%)		
		unmatched	matched w/ EB	matched w/ PSM	unmatched	matched w/ EB	matched w/ PSM
Reduced set of controls							
<i>Maternal characteristics in t_0</i>							
Vocational training ⁺	64.6	68.8	64.5	67.1	-8.9	0.0	-5.3
University ⁺	29.3	15.6	29.3	22.9	33.1	0.0	14.5
Missing uni ⁺	0.5	0.9	0.5	1.6	-3.9	0.0	-10.1
Basic school ⁺	10.4	21.9	10.4	13.7	-31.6	-0.0	-10.2
Intermediate school ⁺	39.9	38.5	39.9	45.6	2.9	0.0	-11.6
Technical college ⁺	6.1	6.3	6.1	4.8	-0.9	0.0	5.5
Highest secondary ⁺	33.1	20.4	33.1	24.1	28.9	0.0	20.0
Other school ⁺	7.0	9.4	7.0	7.0	-8.8	0.0	-0.2
School dropout ⁺	1.7	2.3	1.7	2.6	-4.0	0.0	-5.7
In school ⁺	0.3	0.0	0.0	0.0	7.9	7.9	7.9
School missing ⁺	1.5	1.2	1.8	2.1	2.7	-2.6	-4.4
Migration Background ⁺	19.0	27.5	19.0	20.9	-20.2	0.0	-4.7
Age mother	36.2	36.1	36.2	35.5	1.9	0.2	12.5
<i>Maternal employment history</i>							
Works full-time ⁺	25.4	7.6	25.3	25.6	49.4	0.0	-0.5
Working ⁺	72.5	54.3	72.4	70.2	38.3	0.0	5.1
Actual working-hours	22.3	11.5	22.3	21.9	70.3	0.0	2.2
Desired -actual hours	-6.1	-1.8	-6.1	-4.6	-37.4	-0.0	-11.7
Missing desired ⁺	1.3	1.3	1.3	1.7	0.5	0.0	-3.3
Years part-time	3.1	2.9	3.1	2.9	5.4	0.0	6.0
Years full-time	6.3	5.7	6.3	5.8	11.1	0.0	10.3
Missing LFS-experience ⁺	0.2	0.1	0.2	0.6	4.3	0.0	-6.1
Full-time t-1 ⁺	0.2	0.1	0.2	0.2	40.3	0.0	2.8
Working t-1 ⁺	0.6	0.4	0.6	0.6	29.8	0.0	6.8
Missing working t-1 ⁺	0.2	0.1	0.2	0.1	5.8	0.0	2.2
Working-hours t-1	17.9	9.1	17.9	16.8	56.6	0.0	6.3
Missing working-hours t-1 ⁺	0.4	0.6	0.4	0.4	-30.8	0.0	-6.9
<i>Child characteristics in t_0</i>							
Age child	5.7	5.7	5.7	5.7	0.5	0.3	-3.4
Younger siblings ⁺	40.5	45.9	40.4	42.4	-11.0	0.0	-4.0
Older siblings ⁺	41.9	55.1	41.9	48.2	-26.7	0.0	-12.7
Only-child ⁺	27.2	14.3	27.1	21.7	32.0	0.0	12.7
Female child ⁺	47.6	49.6	47.6	48.5	-4.1	0.0	-1.9
CC none ⁺	60.7	61.3	60.7	60.2	-1.1	0.0	1.0
CC relatives ⁺	28.2	25.4	28.2	30.0	6.3	0.0	-4.1
CC friends ⁺	7.0	4.4	7.0	6.0	11.2	0.0	3.9
CC paid carer ⁺	4.5	2.7	4.5	3.1	9.9	0.0	7.3
Missing CC ⁺	4.0	9.2	4.0	4.9	-21.0	0.0	-4.4
ECEC Hours	4.7	2.0	4.7	4.6	86.2	0.0	3.5
Missing ECEC-hours ⁺	0.8	1.1	0.8	0.8	-3.7	0.0	-0.2
ECEC full-time ⁺	22.9	11.7	22.9	22.7	30.1	0.0	0.4
<i>Partner information in t_0</i>							
Vocational training ⁺	0.5	0.6	0.5	0.5	-18.7	0.0	-3.3
University ⁺	0.2	0.2	0.2	0.2	1.3	0.0	8.6
Missing uni ⁺	0.0	0.0	0.0	0.0	-7.1	0.0	-4.2
Basic school ⁺	28.0	15.9	28.0	24.5	29.5	-0.0	7.9
Intermediate school ⁺	11.8	24.9	11.8	10.6	-34.3	0.0	3.7
Technical college ⁺	25.4	20.5	25.4	32.6	11.8	0.0	-15.9
Highest secondary ⁺	4.1	7.1	4.1	3.8	-13.1	0.0	1.3
Other school ⁺	22.6	19.5	22.6	18.7	7.8	0.0	9.7
School dropout ⁺	6.1	9.0	6.1	7.1	-10.9	0.0	-4.0
In school ⁺	1.3	2.2	1.3	2.0	-7.0	0.0	-6.0
School missing ⁺	0.7	1.1	0.7	0.6	-4.0	0.0	1.1
Migration Background ⁺	0.2	0.2	0.2	0.2	-19.8	0.0	0.8
Age	28.1	32.9	28.1	29.0	-28.2	0.0	-4.8
Working ⁺	0.7	0.8	0.7	0.7	-27.3	0.0	-7.3
Desired -actual hours	-9.9	-9.9	-9.9	-9.4	-0.0	-0.0	-2.5
Missing desired ⁺	0.0	0.0	0.0	0.0	-4.7	0.0	-3.4

Variable	Means treated	Means controls			Standard. Bias (%)		
		unmatched	matched w/ EB	matched w/ PSM	unmatched	matched w/ EB	matched w/ PSM
Actual working-hours	29.2	34.4	29.2	30.3	-24.0	0.0	-5.1
Missing hours ⁺	0.0	0.0	0.0	0.0	-7.1	0.0	0.6
Household characteristics in t₀							
No spouse ⁺	21.4	9.9	21.4	19.2	32.0	0.0	5.3
Home owner ⁺	40.9	55.8	40.9	43.5	-30.1	0.0	-5.2
HH income (in 1000)	48.1	48.6	48.1	44.5	-1.3	0.0	8.5
Village ⁺	7.4	8.0	7.4	11.5	-2.2	0.0	-13.9
Small town ⁺	8.2	12.2	8.2	7.7	-13.2	0.0	2.1
Medium town ⁺	22.1	31.4	22.1	23.9	-21.1	0.0	-4.1
Large town ⁺	17.5	18.4	17.5	22.5	-2.2	0.0	-12.3
Small city ⁺	7.0	9.8	7.0	7.2	-10.3	0.0	-1.0
Medium city ⁺	17.8	12.6	17.8	12.8	14.4	0.0	13.8
Large city ⁺	20.0	7.6	20.0	14.5	36.6	-0.0	14.5
State GDP/1000	233.2	299.4	233.1	210.8	-35.6	0.0	11.0
Unemployment share	9.9	7.8	9.9	10.3	56.5	0.1	-9.5
Additional variables of full set of controls							
Labour market “unobservables” in t₀							
SW definitely not ⁺	2.6	8.4	2.6	3.2	-25.7	0.0	-3.5
SW improbable ⁺	1.7	6.3	1.7	2.6	-23.3	-0.2	-5.8
SW probable ⁺	6.7	15.2	6.7	7.7	-27.6	0.0	-4.2
SW definitely ⁺	15.8	15.2	15.8	16.0	1.6	0.0	-0.5
SW missing ⁺	0.8	0.6	0.8	0.4	2.1	0.0	5.5
SW asap ⁺	8.5	6.3	8.5	9.4	8.7	0.0	-3.0
SW this year ⁺	7.7	8.2	7.7	7.8	-1.8	0.0	-0.3
SW 2-5 years ⁺	6.1	15.4	6.1	7.1	-30.2	0.0	-4.1
SW 5+ years ⁺	1.7	6.4	1.7	1.9	-23.9	0.0	-1.6
SW missing time ⁺	0.2	0.5	0.2	0.1	-5.6	-0.5	1.3
SW full-time ⁺	3.8	2.0	3.8	3.6	10.4	-0.1	0.6
SW part-time ⁺	14.9	29.7	14.9	16.5	-36.2	0.0	-4.6
SW both ⁺	5.0	2.7	5.0	5.5	12.1	0.0	-2.0
SW dont know ⁺	0.5	2.3	0.5	0.7	-14.5	0.0	-1.3
FJ not applicable ⁺	75.0	60.0	75.0	72.7	32.6	0.0	5.3
FJ easy ⁺	3.3	7.1	3.3	3.1	-17.3	-0.1	1.2
FJ difficult ⁺	14.9	23.8	14.9	17.2	-22.6	0.0	-6.3
FJ almost impossible ⁺	6.5	8.5	6.5	6.7	-7.5	0.0	-0.7
FJ missing ⁺	0.2	0.6	0.2	0.3	-5.7	0.0	-0.9
Job-search yes ⁺	7.7	6.5	7.7	8.6	4.6	0.0	-3.5
Job-search no ⁺	16.5	30.0	16.5	17.7	-32.4	0.0	-3.1
Job-search missing ⁺	0.0	0.1	0.0	0.0	-5.2	-2.1	-1.7
SI yes ⁺	11.5	13.1	11.5	13.3	-5.0	0.0	-5.6
SI no ⁺	12.5	23.2	12.5	12.9	-28.0	0.0	-1.0
SI missing ⁺	0.2	0.4	0.2	0.1	-4.2	-0.5	1.3
N	1,278	2,976					

Notes: EB=entropy balancing; PSM=propensity score matching; SW=searching for work; FJ=finding a job; SI=starting job immediately; CC=childcare. Summary statistics for treated, all controls and matched controls (indicators for state, year and SOEP sample not shown). The first two columns present the variable means before matching for treated and controls. The third and fourth column show the means for the re-weighted control group according to entropy balancing (EB) and kernel matching, a propensity score method. The last three columns display a measure for the quality of the matching process. The standardized bias is defined for each conditioning variable s as $SB_s = 100 \cdot \frac{\bar{s}_1 - \bar{s}_0}{\sqrt{\frac{1}{2}(\sigma_{s1}^2 + \sigma_{s0}^2)}}$, where \bar{s}_1

and \bar{s}_0 are the means of treated and controls, respectively, and σ_{s1}^2 and σ_{s0}^2 the corresponding variances. ⁺ indicates that the mean represents a percentage share.

Source: SOEP v31.

Table A.2: Tobit and hurdle regression

	Main	Tobit	Hurdle
Working	0.075*** (0.024)		0.074** (0.036)
N	4254		4254
Hours	2.779*** (0.836)	2.894*** (0.894)	2.841*** (1.074)
N	4254	4254	4254

Notes: The table presents marginal effects from the main model (column 1), Tobit regression (column 2), and Cragg hurdle regression (column 3) for the full sample. Source: SOEP v31. Robust standard errors clustered at the mothers' level in parentheses, significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.