As in many cases: it depends. While the availability of all-day schools does significantly increase the probability that children are in full-day childcare (by 7.6 to 17.6 percentage points at the mean all-day school supply), overall no effect on labor market participation can be found. If, however, all-day school supply is sufficiently high (over 20% on average) mothers who have their children in full-day care due to the presence of all-day schools increase their labor supply by 13.7 hours per week. These results are obtained exploiting the expansion of all-day school supply in Germany which was rolled out since 2003 in a staggered fashion across states and counties, thus creating quasi-experimental variation in the childcare choices, and implicitly prices, faced by mothers. Using the county-level share of all-day primary schools as an instrument for children being in full-day care this paper estimates a local average treatment effect (LATE) on maternal labor supply. Further, utilizing data on county-level take-up rates allows to rule out potential biases through otherwise unobserved parental preferences. Compared to existing work this paper contributes by employing the to date most comprehensive and precise dataset on county-level all-day school supply as well as by exploring the conditions under which they show an effect.

JEL: J22, J13, D04
Keywords: Maternal labor supply, public childcare, all-day schools

* ZEW Mannheim, Germany. Questions and comments are welcome via email to lukas.riedel@zew.de.

I thank Christina Gathmann (who in addition supported the collection of the data), Sebastian Siegloch and Holger Stichnoth as well as seminar participants at ZEW and Compié 2018.

The data used in this paper were made available by the German Socio-Economic Panel Study (SOEP) at the German Institute for Economic Research (DIW), Berlin.
1 Introduction

This paper analyzes the effect of a large expansion of all-day primary schools on maternal labor supply in Germany.

Maternal labor supply and its driving factors is a long standing topic in empirical economic research (Heckman and Killingsworth 1986 give an overview of early approaches). In this field costs and availability of childcare is often viewed as a key determinant in mothers’ decision making (see for instance Blau and Robins 1988, Kimmel 1998 or Jaumotte 2003).

Recent research often focuses on childcare for young children before school entry. In most cases policy reforms which expanded the availability, i.e. decreased the price of childcare are studied. Studies from outside Germany find different effects, ranging from a strong positive (Barker, Gruber, and Milligan 2008) to no (Havnes and Mogstad 2011) association between childcare and maternal labor supply. Also for Germany predominantly expansions of childcare are studied as in recent years the public provision of childcare was increased strongly in response to ongoing demands of parents (Spieß, Berger, and Groh-Samberg 2008). Main subjects of analysis are two policy reforms that introduced a legal claim to a place in childcare for 3 to 6 year olds in 1996 and an expansion of this claim to 1 to 3 year olds in 2013. Wrohlich (2005) estimates the size of the excess demand for childcare. Studies by Bauernschuster and Schlotter (2015), Wrohlich (2006) or Coneus, Goeggel, and Muehler (2009) find positive effects of childcare on labor supply, while Busse and Gathmann (2015), Bauernschuster, Hiener, and Rainer (2016) and Bauernschuster and Bork (2016) add findings on children’s outcomes, divorce rates and fertility. Gathmann and Sass (2018) on the contrary investigate the opposite case of an increase in the price of childcare that is associated with decreasing maternal labor supply. General considerations on the importance of childcare are added by Boll (2010) who quantifies the financial loss for mothers who stay out of the labormarket.

There is overall fewer research that considers childcare for children of primary school age. Lundin, Mörk, and Öckert (2008) analyze a reduction of childcare prices for one to nine year olds in Sweden using a difference-in-differences approach and find no significant effects on maternal labor supply. Contreras, Sepúlveda, and Cabrera (2010) and Berthelon, Kruger, and Oyarzün (2015) estimate panel data models for a reform in Chile using the regional availability of all-day schools as explanatory variable. They find positive and significant effects. Felfe, Lechner, and Thiemann (2013) exploit spatial discontinuity in childcare availability along cantonal borders in Switzerland finding a positive effect as well.

In Germany a reform that initiated the large-scale move from half-day to all-day school decided by the federal government in 2003 provides a possible object of research. As matters of education policy are usually decided on the state-level this reform offers a rare opportunity to study nationwide how mothers of primary school aged children react to an increased supply with childcare possibilities.

There are key characteristics of the German labor market which suggest the presence of a significant demand for childcare for this specific age group. Figure 1 plots mothers’ labor supply against their children’s age. They increase their hours worked as their children get older, the growth is particularly strong for children younger than five indicating that the reforms targeted at young children can considered to be successful. Also for children older than ten the trend is increasing, arguably an effect of the declining demand for care as children become more independent. However, in the age range from five to ten mothers exhibit almost no increase in their labor supply, i.e. especially at the time when children visit primary school.

To analyze if this potential demand for care translates to labor supply as the availability of all-day schools increases this paper uses the share of all-day schools in all primary schools in a county (Landkreis) as an instrument for a child being in full-day care. With this instrument a local average treatment effect of all-day school supply on maternal labor supply is estimated. Data on all-day schools are collected from various official
sources, household-level data is provided by the German Socio-Economic Panel (SOEP). The main findings suggest that all-day schools contribute significantly to full-day childcare of primary school aged children. However, a significant effect on labormarket participation of mothers cannot be found. In federal states where the supply with all-day schools is high an effect of 13.7 additional working hours per week for mothers who are already in the labormarket and have their children in full-day care due to the availability of all-day schools can be identified. These results are overall robust and consistent with specifics of the German all-day schools system. Further, an impact of otherwise unobservable parental preferences can be ruled out by using data on all-day school utilization.

After their expansion started all-day schools in Germany were first mostly studied regarding their effects on children’s educational outcomes (see Bundesministerium für Bildung und Forschung 2012 or StEG-Konsortium 2010). This object of research is, however, not entirely suitable for research that considers Germany as a whole or at least large parts of it as there are significant differences between federal states such that the major feature German all-day schools have in common is the provision of extra care. For the effects of this additional time of childcare Beblo, Lauer, and Wrohlich (2005) provide an ex-ante microsimulation that suggests moderate increases in labor supply by 1 percentage point (pp) on the extensive margin and 4 pp on the intensive. Tobisch (2013) and Nemitz (2015) mostly rely on the comparison of mothers who utilize all-day schools with those who do not while others exploit the quasi-experimental nature of the expansion of all-day schools. Shure (2016) uses a difference-in-differences approach on data from Bavaria, Hesse, Rhineland-Palatinate and Schleswig-Holstein where all-day school supply is coded binary finding an effect of around 7 pp on the extensive margin but no impact on hours worked. The approach of Dehos and Paul (2016) is, similar to the one of this paper, to estimate an instrumental variable model with the county-level share of newly created all-day schools as instrument. They find no significant effects, neither on the intensive nor on the extensive margin. Similar to Shure’s work Boll and Hoffmann (2017) code all-day schools binary to estimate the direct effect of their county-level availability on la-
bor supply of all mothers in a county when their child is at the age of 15. They find an effect of 2.8 additional weekly working hours for the presence of an all-day offer. Complementary, they find some indication that the use of all-day schools can be beneficial for children’s school performance. Gambaro, Marcus, and Peter (2016) approach the potential selection bias of working mothers into all-day school utilization with a matching strategy. Their findings suggest an increase by 2.6 weekly working hours and a 11.4 pp higher employment probability. These results, however, combine mothers who utilize all-day schools and those who have their children in other childcare facilities, thus neglecting potentially different effects.

This paper’s main contribution is the use of the to date most comprehensive and precise dataset on county-level all-day school supply. It points out that not the sole presence of all-day schools but rather their coverage is crucial for their effect on labor supply. Further, this is the, so far, only work that uses data on utilization rates to control for impacts of otherwise unobserved parental preferences.

The remainder of this paper is organized as follows. Section 2 gives an overview of the important aspects of maternal labor supply and childcare in Germany. Section 3 presents the estimation framework as well as considerations on the required assumptions and if they hold. The data is introduced in section 4. Estimation results and robustness checks are presented in sections 5 and 6 before section 7 concludes with a discussion and some remarks.

2 Institutional Background

Maternal labor supply in Germany increased over the last years. As figure 2 indicates the growth of the labormarket participation rate between 2006 and 2014 is among the highest compared to other European countries. If the current levels are considered Germany ranges slightly above the averages of OECD as well as EU (OECD 2016). However, in comparison to other countries in Central Europe fewer of the German mothers work. Additionally, the participation rate (according to OECD 2017 69% in 2013 and 66.4% in 2012 according to Bundesministerium für Familie, Senioren, Frauen und Jugend 2014) omits the fact that most working mothers only work less than 32 hours per week (Bundesministerium für Familie, Senioren, Frauen und Jugend 2014).

As this paper discusses all-day primary schools labor supply of mothers with children in primary school is of particular interest. While mothers increase their participation rates when their children get older this increase slows down around the age of school entry (Bundesministerium für Familie, Senioren, Frauen und Jugend 2014). Similar results can be found at the intensive margin. Considering data from the larger, and thus more representative, Microcensus dataset (figure 8 in the appendix) supports the findings on the sample used in this paper (depicted in figure 1 in the previous section). Until children reach school age at around six years their mothers increase their hours worked more strongly than while they are in primary school. With the end of primary school at ages around ten the increase gets slightly stronger again.

Besides other reasons, for instance the German system of income tax splitting for married couples which imposes potentially adverse incentives on secondary earners, the availability of childcare is a crucial element in labor supply decisions of mothers. Similar to labor supply the utilization of childcare shows an increasing trend. For children under the age of three childcare was expanded over the past years, since 2013 parents have the legal claim to a place for one to three year olds (Sozialgesetzbuch Achtes Buch – Kinder- und Jugendhilfe (SGB VIII) 2012). While in 2006 13.6% of them where in care, this share increased to 32.3% in 2014 with an increasing share of full-day care (Bundesministerium für Familie, Senioren, Frauen und Jugend 2015). Regardless of this increase parents still indicate excess demand for childcare for under three year olds (Bundesministerium für Familie, Senioren, Frauen und Jugend 2015). For children aged three until school entry such an excess of demand is unlikely as the legal claim to a place in Kinder-
Figure 2: Comparison of labor market participation rates of mothers (aged 15–64) with at least one child under the age of 15 in different European countries for the years 2006–2014 (data on Germany only available until 2013). Illustration by author, based on OECD (2017).

The "Garten" came into effect in 1996 (Spieß, Berger, and Groh-Samberg 2008), to date over 95% are in care. For both age groups over 80% of those in care are there for 25 hours per week or longer, about one half more than 35 weekly hours. Compared to publicly subsidized childcare private caregivers play a negligible role (Spieß, Berger, and Groh-Samberg 2008).

When children start visiting primary school care can be supplied by either all-day schools or other childcare facilities (in German Horte) where the latter are established for a longer time than all-day schools. The utilization of these facilities increased from 10.6% in 2006 to 16.2% in 2015. For all-day schools this rate is currently about twice as high (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland 2016). The extent to which Horte are utilized differs across federal states. They are the predominant form of childcare for children in primary school in those East German states that are not considered in this paper. In other states responsibility for childcare was transferred completely to all-day schools in the course of their expansion (Bundesministerium für Familie, Senioren, Frauen und Jugend 2015).

This expansion of all-day schools did not only happen in some states but nationwide. A factor that is notable as Germany’s political organization is that of a federal republic. While the federal government in general has most of the legislative competence with the 16 states fulfilling mostly administrative functions matters of education policy are an exception. The legislative framework as well as the administration of schools is almost entirely provided by the states with the consequence of very heterogeneous educational environments in Germany. The expansion of all-day schools, however, was primarily financed by the federal government, specifically by the "Investment Program Future Education and Care“¹ (Bundesministerium für Bildung und Forschung 2003). Between 2003 and 2009 (Bundesministerium für Bildung und Forschung 2006) an amount of €4 billion was supplied by the federal government along with additional €400 million by the states. Given that the federal

¹Translation of the German term "Investitionsprogramm Zukunft Bildung und Betreuung"
government’s total spending for investments in 2003 amounted to around €26bn (Bundesministerium der Finanzen 2009) this program can be considered important. Of the around 7200 schools which received funding 52% were primary schools (Bundesministerium für Bildung und Forschung 2009).

There were various motivations for the investment program. While the wishes of parents as well as the intent to provide additional means of childcare did play a role, arguably, the most prominent aspect in public debate was that all-day schools were viewed as capable of improving the quality of education (Bundesministerium für Bildung und Forschung 2017). The weak performance of German students in international comparisons was a widely discussed issue at that time what might explain that such a large program was funded by the federal government.

The results of the program are evident. In 2002 10.3% of all primary schools in Germany had an all-day offer and 4.2% of students used one (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland 2008). Until 2014 the share of schools increased to 53.3% with 33.1% of students visiting an all-day school (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland 2016).

As of today the way all-day schools are implemented is very heterogeneous across states. The definitions of features an all-day school needs to fulfill vary in terms of organization, pedagogical objectives and learning contents, even the aim of supporting families by providing means of childcare is not a shared characteristic of states’ legislation (Berkemeyer 2015). Nevertheless, a, if not the, common feature of all-day schools is that they supply childcare in addition to the time usually spent in school. Minimum standards holding across all federal states are defined by the Standing Conference of the Ministers of Education and Cultural Affairs of the Länder. These standards require that a school has to provide an all-day offer of at least seven hours on at least three days per week to qualify as an all-day school (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland 2016). The amount of time offered beyond this requirement, again, varies across states. Klemm and Zorn (2016) estimate the time children are in school by reviewing the current legal situation. In addition to the usual time in class it ranges from 7.9 hours per week in Thuringia to 22.4 hours/week in Hesse (the results for all states are given in figure 9 in the appendix) with a mean of 13.7 hours/week². In this estimation fulfilling the minimum standards would relate to additional time of around 8.5 hours per week. Moreover, it is of note that only in Bremen, Hamburg, Hesse and Schleswig-Holstein all-day schools are open for all five weekdays; most states offer four days of care, North Rhine-Westphalia, Saxony and Thuringia on three. This highlights that, despite their name, all-day schools do not necessarily offer childcare for the full day on every workday.

Another common feature is that usually three different types of all-day schools are offered (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland 2016). In completely integrated all-day schools all students have to participate in the all-day offer. The two other types give parents more freedom of choice as in the partially integrated type only a part of the students (e.g. some classes) are required to participate and in the open type of all-day schools participation is voluntary for each student.

As it simplified the data collection this paper uses data on all types summed up without further differentiation. The completely integrated type could pose a threat to the estimation as, compared to voluntary, forced participation arguably has different consequences for mothers’ labor supply decisions. This case, however, is unlikely. In 2014 only 3.9% of all German all-day schools were of the completely integrated type³. Their share in this paper’s sample will be even smaller since the state with the highest supply of them (Bremen) is not included in the dataset.

---

² Even though incorporating these information in the estimation can provide additional insights this paper does not use them as they are based on the current state of legislation that cannot necessarily be generalized to previous years.

³ Own computation based on Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (2016)
3 Empirical Strategy

This paper uses the supply with all-day primary schools as an instrument for full-day childcare to identify their effect on maternal labor supply in an instrumental variable (IV) strategy. The approach to estimate the effect of all-day schools with the intermediate step of considering full-day childcare has the advantage of being able to disentangle different single effect. The two-step procedure of IV estimates in the first stage how the supply with all-day schools affects the probability of a child being in full-day care. In the second stage this estimated probability is used to compute how mothers change their labor supply if their children are in an all-day school. Further, reduced form regressions of maternal labor supply on all-day school availability provide a measure of their effect on all mothers without taking their children’s care status into account.

In the following this section discusses the applicability of the IV method as well as the effects it estimates.

To justify the use of IV it has to be noted why a simple OLS regression is likely to fail. For the remainder \( Y_i \) denotes a work outcome, either on the intensive or extensive margin, of the mother of child \( i \), a binary variable \( C_i \) indicates if child \( i \) is in care for the whole day.

If the structural equation

\[
Y_i = C_i + \epsilon_i \tag{1}
\]

is estimated directly two potential problems can arise. First, as a mother’s work decision as well as her choice for childcare are likely to be partly determined by each other, a simultaneity bias can be present. Equation (1) then does not estimate a causal effect of \( C_i \) on \( Y_i \) as only equilibrium outcomes where it is unclear which side of the equation has what impact on the other and vice versa can be observed. Second, also omitted variables can bias the estimation. One can think of a variety of factors which impact mothers’ labor supply decisions. Some of them can be measured and controlled for directly while others are not measurable but can be proxied by others. For instance, nonlabor-income can be quantified whereas information on a mother’s education can only give hints about her preference for her children’s education and in turn her demand for childcare. However, there still remain variables that are unobservable. For example, decisions on maternal work and childcare can be assumed to be, among other factors, shaped by a number of norms and beliefs. Such cultural influence can be inherited from the parents, from growing up in an urban or more rural milieu as well as religion and the beliefs of the partner can play a role. Likewise, of multifaceted combinations of these and other factors can be thought. As these variables in (1) are correlated both with the dependent and the explanatory variable as well as they are, by their nature of being unobservable, part of the error term, \( C_i \) and \( \epsilon_i \) are correlated with the consequence of biased estimation results. Simultaneity as well as omitted variables lead to endogeneity of \( C_i \), a causal effect cannot be identified.

By choosing an appropriate instrument that is not correlated with \( \epsilon_i \) but correlated with the endogenous regressor IV can be used to circumvent these issues. This paper chooses the availability of all-day schools as instrument \( Z_j \). \( Z_j \) measures the share of all-day primary schools in all primary schools in county \( j \) where child \( i \) lives.

3.1 Local Average Treatment Effect

Even though, IV is an appropriate estimation method for the impact of full-day childcare on maternal labor supply a justification for its use in the context of assessing the question of the effect of all-day schools is still to be given. This justification is provided by the concept of local average treatment effects (LATE). First formulated by Imbens and Angrist (1994) it provides a framework for interpreting instrumental variable estimates in the case of heterogeneous potential outcomes.

Consider the general case of a treatment which affects some outcome. This effect is to be estimated by using an instrument that affects the treatment status. From the whole population of interest there might only be a

---

4 This one and the following section build on Imbens and Wooldridge (2007) and Imbens and Angrist (1994)
subgroup that changes its treatment status in reaction to
the instrument (those who comply) while others do not.
Such a situation could, for instance, arise if an instru-
ment rather changes incentives for being treated than
ultimately assigning treatment. As there are subgroups
of the population that react differentially to an instru-
ment potential outcomes for these groups are not the
same but heterogeneous. Solely for the compliant sub-
population a treatment effect can be estimated, not for
those whose treatment status is unaffected by the in-
strument. This estimate is therefore necessarily local
for the compliant subgroup. It cannot be generalized as
the (theoretical) effect for those who not change their
treatment status could be different.

In the present case where the effect of full-day child-
care on mothers’ labor supply is to be estimated by us-
ing the supply with all-day schools as an instrument
mothers’ potential reactions fall in two categories. First,
they either never send their children to full-day care
or they even would do so without the presence of all-
day schools (by utilizing other means of care provision
like Horte or relatives). Their behavior is unaffected by
the instrument. The second possibility⁵ is that they be-
long to the compliant subpopulation and send their chil-
dren to all-day schools, thus in full-day care. As LATE
only uses this second group, the results obtained repre-
sent the effect of a child being in full-day care on their
mother’s labor supply only for those children who are
in care because of the presence of all-day schools in their
county. Those children who are in all-day schools but
would be cared of somewhere else if there were no all-
day schools are not captured.

Even though using an IV strategy requires the addi-
tional step of looking at how the full-day care status of
children changes the considerations on the LATE show
that the estimated effect is still that of all-day schools. If
a different instrument was used, for instance a reform
of tax deductibility of the costs of childcare by private
providers, arguably a different subpopulation would re-
act such that the LATE for this instrument was differ-
ent.

In addition to the specific LATE estimates the reduced
form estimations of the IV procedure provide more
general effects. They estimate how the supply with all-
day schools affects work outcomes of all mothers eligi-
ble, i.e. those who have children in primary school, re-
gardless of the children’s care status. These estimates
relate to an Intention to Treat Effect (ITT).

3.2 Identification

For the LATE framework to be applicable the following
assumptions have to be satisfied. It is considered that
a mother’s work outcome \( Y_i(c, z) \) depends on the the
child’s full-day care status \( C_i = c \) that is either 0 (\( C_{0i} \))
or 1 (\( C_{1i} \)) and the value of the all-day school instrument
\( Z_j = z \).

**Independence** The instrument is, with respect to \( Y_i \) and
\( C_i \), as good as randomly assigned, i.e. the availability
of all-day schools in a county is unrelated to labor and
childcare decisions of mothers in this county:

\[
[ Y_i(c, z); \forall c, z, C_{1i}, C_{0i} ] \perp Z_i. \tag{2}
\]

Reduced form estimates and first stage can therefore
be interpreted as causal effects. There is the possibility
that there are factors specific to the state as well as the
county level which have an impact on where how much
all-day schools were established. On the state-level this
refers to the differences in all-day school supply dis-
cussed in section 4.2 and on the county-level some states
might have targeted the extension of all-day schools to
counties with, for instance, higher unemployment. Be-
cause of this possibility appropriate control variables
are added such that independence can be assumed to
hold conditional on these controls.

**Exclusion** For the exclusion restriction to hold the in-
strument needs to have its impact only through a single
causal channel. All-day schools have their impact on if
and how much mothers work solely through the indi-
rect channel of providing additional means of childcare.

---

⁵ Summarizing potential reactions to treatment in two groups is a
simplification that holds as long as the monotonicity assumption,
to be discussed below, holds.
As long as a child’s care status stays constant a (hypothetical) change in the value of the instrument does not affect the work outcome:

\[ Y_i(c, z) = Y_i(c, z') \quad \forall \ c, z, z'. \quad [3] \]

If there was a direct effect of all-day schools on maternal labor supply it would have been necessary to include \( Z_j \) in the initial regression of \( Y_i \) on \( C_i \) and thus by using it as instrument rather than as regressor one would re-introduce the omitted-variable bias which is to be circumvented by using an instrumental variable estimation. Besides the technical reasoning the exclusion restriction highlights that the LATE estimation, by requiring the instrument to only act through one channel, computes an effect that can solely be attributed to all-day schools. If the exclusion restriction is satisfied cannot be tested directly (at least in the present case of an estimation with one instrument), however section 6 presents two approaches to test if all-day schools either have another impact by providing half-day care or if their effect is through employing mothers as teachers or caregivers.

**First Stage** Together with independence and exclusion this is an assumption known from IV estimation. It ensures that the instrument has indeed an impact on the endogenous regressor, i.e. that a change in the all-day school supply is associated with a change in children’s care status:

\[ E[C_i | Z_i = z] - E[C_i | Z_i = z'] \neq 0 \quad \text{for} \ z \neq z'. \quad [4] \]

**Monotonicity** This assumption is special to LATE. It requires that, even though the instrument does not need to have an effect on the whole population, members of the compliant subpopulation are affected in the same way\(^6\). Specifically, this means the expansion of all-day schools only increases the probability that a child is in full-day care, it must not decrease it. As all-day schools are an optional offer to parents that adds to an existing set of choices but does not restrict it, this assumption seems satisfied. For the compliant subpopulation a violation of monotonicity would mean that there are parents who would send their children to full-day care if there was no all-day school offer in their county and would not do so if their was, or, more general, that with increasing availability of all-day schools there are parents who more often make the decision against sending their children there. Given the increase in supply with childcare possibilities is associated with a decrease of its price (both direct and in terms of opportunity costs) this would correspond to childcare being a Giffen good. Intuitively, such preferences seem very unlikely. However, there is still the possibility that a growing number of all-day schools crowded out other childcare facilities such that eventually the number of available places decreased. To mitigate this concern one can look at data on childcare facilities. The data\(^7\) is in a consistent manner only available for the years 2007 and later and, additionally, it depicts facilities for children of a broader age-range (up to 14) so it will not be used in the later analysis. It is still suitable to provide an impression if there is crowding out. Four different indicators for the supply with childcare facilities at the county level are considered. These are the overall number of childcare facilities, the number of facilities for school-aged children (five to 14 years old), the overall number of places in facilities and the number of persons employed there. They are analyzed in regressions models with all-day schools as explanatory variable of interest. Fixed-effects estimators are used to account for time-persistent heterogeneity across counties. Additionally, it is allowed for a general as well as for state-specific time trends as states extended all-day schools in a differential fashion. As control variables a county’s per capita GDP and unemployment rate proxy for its economic situation, population density distinguishes urban and rural areas and the share of under-ten year-olds accounts for the overall demand for childcare. In the case that the assumption of crowding-out was true one would expect a significantly negative effect of all-day schools on other child-

---

\(^6\) It has to be noted that this assumption as well as the concept have been subject of criticism. Heckman (2005) identifies it as an asymmetry in IV literature – "Outcomes are permitted to be heterogeneous among persons in a general way. Choices of treatment are not permitted to be heterogeneous in a general way." – however admits that LATE estimates serve their purpose in evaluating policies and produce "economically interpretable parameter[3]" (Heckman 1997).

\(^7\) See also section 11.3 in the appendix.
where some of the controls show significant impacts, especially the unemployment rate, population density and the share of children, the share of all-day schools does not. Even though, three of four co-efficients are negative the smallest p-value is 18.3% (detailed results can be found in table 7 in the appendix). Therefore, the assumption that a larger supply with all-day schools crowds out other childcare providers can be rejected and thus, the LATE assumption of monotonicity seems, as far as it can be tested, fulfilled.

A formal definition of the monotonicity assumption requires to note that the instrument is not a binary variable as assumed in most representations of the LATE. At a global level, i.e. when all counties together are looked at, the share of all-day schools can be viewed as continuous in the range from zero to one. However, at the county level, where the instrument exhibits its effect, it takes on at most 13 distinct values as the years 2002 to 2014 are observed. Thus, treadng the instrument as multivalued rather than continuous is plausible and simplifies the formal representation.

Its multivalued nature is now represented by denoting it as a function \(g(.)\) of \(Z_i\) where \(Z_i\) has support over \(z_0, …, z_K\) with \(K \leq 12\). Monotonicity requires that the instrument values are ordered such that for any change from \(z_{k-1}\) to \(z_k\) where \(z_{k-1} < z_k\) it holds that

\[ P(z_{k-1}) < P(z_k) \quad \forall k \]

Where \(P(z)\) abbreviates the expectation of the full-day care indicator \(C_i\) given an instrument value \(z\), i.e. \(P(z) = E[C_i|Z_i = z]\). That means, for any increase in the share of all-day schools the probability of a child being in full-day care does not decrease.

Given all assumptions hold the IV estimator as the fraction of the covariance of the instrument and the outcome of interest and the covariance of the instrument and the endogenous regressor is given as

\[ \hat{\beta}_{IV} = \frac{\text{Cov} \left( Y_i, g(Z_i) \right)}{\text{Cov} \left( C_i, g(Z_i) \right)} = \sum_{k=1}^{K} \lambda_k \beta_{z_k,z_{k-1}}. \]

This is a weighted sum of \(K\) (theoretical) IV estimations in which only the instrument pair \(z_k\) and \(z_{k-1}\) is considered instead of the full range of available values. \(\lambda_k\) assigns weights to the pairwise estimates.

\[ \lambda_k = \frac{(P(z_k) - P(z_{k-1})) \sum_{l=1}^{K} \pi_l (g(z_l) - E[g(Z)])}{\sum_{m=1}^{K} (P(z_m) - P(z_{m-1})) \sum_{l=m}^{K} \pi_l (g(z_l) - E[g(Z)])} \]

with \(\pi_k = Pr(Z = z_k)\) denoting the probability mass function of \(Z_i\). The weights \(\lambda_k\) are nonnegative (by the monotonicity assumption that \(P(z_{k-1}) < P(z_k)\)) and they sum up to one. They further depend positively on the expected size of the change in \(C_i\) associated with a change of \(Z_i\). The function \(g(.)\) is based on the first stage estimations of how \(Z_i\) impacts \(C_i\).

In the case of a multivalued instrument the LATE estimate is a weighted average of single estimates for any change in the value of the instrument. Instrument values that appear more often in the distribution of \(Z_i\) are given a larger weight. The single estimates can be different, thus the model allows for the possibility that an increase in all-day school supply from 0 to 5% has a smaller impact on the probability that a child is in full-day care and in turn on their mother’s labor supply than a change from 60 to 80%.

### 3.3 Estimation Setup

Taking everything discussed above together yields the following different specifications of the estimation equation

\[ Y_{it} = \beta \hat{C}_{it} + D_{it} + D_{i} + \epsilon_{idt} \]

\[ = \beta Z_{it} \tilde{Y}_{it} + D_{it} + D_{i} + \epsilon_{idt} \]

\[ Y_{it} = \beta \hat{C}_{it} + X_{it} + D_{it} + D_{i} + \epsilon_{idt} \]

\[ = \beta Z_{it} \tilde{Y}_{it} + X_{it} + D_{it} + D_{i} + \epsilon_{idt} \]

and

\[ Y_{it} = \beta \hat{C}_{it} + X_{it} + W_{ijt} + D_{it} + D_{i} + \epsilon_{ijt} \]

\[ = \beta Z_{ijt} \tilde{Y}_{it} + X_{ijt} + W_{ijt} + D_{ijt} + D_{i} + \epsilon_{ijt} \]

The estimated coefficient of interest, \(\hat{\beta}\), gives the effect of a child being in full-day care on their mother’s
work outcome $Y_{it}$ conditional on covariates. It is obtained by estimating $[6]$ in a two-stage-least-squares model where $\gamma$ is the effect in the first stage in which $C_{it}$ is to be explained by the share of all-day schools in the county $j$ where $i$ lives ($Z_{jt}$).

Compared to the simplified form used before the equations incorporate, first, the fact that observations for several years are used, indicated by the time index $t$, and, second, a set of control variables.

$[8]$ does not include covariates, only fixed effects for federal states ($D_s$) as the level on which matters of education policy are usually decided.

$[9]$ adds covariates that control for characteristics of the child, respectively their mother and the household they live in ($X_{it}$). Specifically these are the child’s and mother’s age, an indicator if the mother is a single parent, her hourly wage, indicators for university-level or no education (such that those owning a vocational degree are the comparison group), nonlabor income (including the income of the husband, if present), an indicator for a migration background and the number of younger and older children in the household, ordered in age groups (babies, infants and those above primary school age). Together these covariates depict the most important determining factors a mother faces when making work decisions.

The full specification $[10]$ additionally takes the county-level into account by using the unemployment rate, per capita GDP (to depict a county’s economic situation and labormarket conditions) and the population density (to distinguish urban an rural areas) as control variables ($W_{jt}$). Further, the state fixed effects are replace with fixed effects for counties ($D_j$) to account for differences that are not picked up by the already included variables, for instance differences in policy making or attitudes of the population.

All models include year fixed effects ($D_t$) to control for time trends.

Since the supply with all-day schools is measured per county standard errors of observations in a county can be correlated. They are therefore clustered at the county-level.

This setup enables to estimate models which incorporate factors that have their effect at different levels. Thus, potentially different results for different models can be easier related to their origin.

4 Data

This paper uses data from two different levels of observation. Information about households is supplied by the German Socio-Economic Panel (SOEP). On the county-level self-collected data on all-day schools is used together with other county characteristics.

4.1 Household-Level Data from the SOEP

The SOEP is a survey among German households conducted annually since 1984. It contains a variety of information, from basic socio-economic characteristics to personal beliefs and psychological traits. In about 12 000 households all household members of at least the age of 17 are questioned, information on children is usually collected through the parents. Even though there have been several additions to the sample the SOEP aims to survey the initial households from 1984 and their descendants as well as to be representative for the “adult population living in private households” (Wagner, Frick, and Schupp 2007, p. 10). If a lower administrative level than Germany as a whole is considered the SOEP-dataset only provides representative information for some large federal states (Baden-Württemberg, Bavaria, North Rhine-Westphalia) (Goebel 2017). Therefore, as analyzing single states might be a threat to external validity, only observations form groups of states are used.

In this paper the SOEP waves from 2002 to 2014 are used. From the original data an estimation sample of 24 663 children in 17 696 households is constructed. In
the construction process several restrictions were applied. Most important, that the analysis is restricted to children of primary school age. This is done for two reasons where the first one is comparability. Primary schools in Germany are a type of school that is common for all federal states and obligatory to visit. With a child’s entry to secondary school parents can choose between at least two types of schools (with a number of subtypes in some states) which are designed differently across federal states and show different supply with all-day schools. In comparison, primary schools are common to all states, obligatory to attend and the only type of school for children at this age range, so they constitute a relatively homogeneous group. Second, the potential considerations of mothers in terms of childcare and labor supply have to be taken into account. When a child enters primary school at the age of usually five or six years a certain amount of care is offered per se whereas any kind of childcare before school entry is utilized on a voluntary basis. Thus an incentive to consider either re-entering the labor market or to increase hours worked is given. For mothers of children in secondary school the consideration to re-enter the labor market seems less likely as at that point they already stayed out of it for about ten years thus taking up work again will be more difficult. In addition, as they get older children in secondary school become more independent such that their parents need to rely on opportunities for care to a smaller extent when making work decisions. Thus children and their treatment at schools as well as their parents and their decision making can assumed to be more consistent as long only primary school children are considered.

10 In 2014 69.6% of "Hauptschulen" (secondary schools for the years 5 to 9) had an all-day offer in comparison to only 52.4% of "Realschulen" (secondary schools for the years 5 to 10) (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland 2016)

11 This setup requires the assumption that mothers are not constrained in their decision through demand. This assumption can be justified by using control variables to account for characteristics of the mother and the county she lives in as well as by noting that the group of mothers of primary school students in Germany is small compared to the overall workforce, thus, general equilibrium effects of the expansion of all-day schools are unlikely.

Table 1: Summary Statistics of the SOEP data on children and mothers/households. Additional summary statistics are listed in table 8 in the appendix.
Instead of focusing on the youngest child (as e.g. Dehos and Paul 2016 do) this paper also includes children who have younger siblings. As care possibilities for younger children were expanded potentially all mothers with children in primary school benefit from all-day schools. Mothers might still be more reluctant to give younger children in care and their presence can have an effect on how the household organizes childcare, so for their number as well as for the number of older children in a household is controlled.

Even though the SOEP includes an indicator if a child attends primary school it is not used. That is, because it, first, contains a relevant number of missing values, and, second, nonmissing values contain the bias that over 80% of households are questioned in the first half of the year¹² while school starts in autumn. Six is the age of school entry intended by legislation, however, pre- or postponing is possible depending on parents' preferences and state-specific regulations (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland 2015). Therefore, five is assumed to be the earliest age for school entry and seven the latest with the latter implying school attendance up to the age of eleven. The robustness of this assumption is tested in section 6 by using a narrower age range.

For all children in the sample binary indicators if they are in care for the full or half day are constructed. As the SOEP-data do not allow to consistently identify if a child visits an all-day school (Marcus, Nemitz, and Spieß 2013) it is possible that the care-indicators also capture children who are cared for in a Hort or some other facility. This is no threat to the validity of the estimation as the mean of this indicator is 15% whereas for the same period of time (2002–2014) on average 18.3% of primary school students participated in an all-day offer¹³. Therefore, first stage estimations of the impact of all-day schools on children being in full-day care may rather constitute a lower boundary.

The outcome of interest, maternal labor supply, is analyzed both at the intensive and the extensive margin. Hours of work are measured per week, the dummy variable for labormarket participation equals to one if a woman works either full- or parttime. A further differentiation, for instance testing if mothers switch from part- to fulltime jobs, is not feasible as not enough observations are available. Mothers in vocational training or education are not considered since their labor supply can be assumed to be, at least, partly fixed. To account for outliers, for instance women who report they do work and zero weekly hours at the same time, observations in the largest and smallest percentile of the distribution of working hours are dropped.

In addition, personal characteristics of the mothers depicting their socio-economic status are collected as control variables (see section 3.3). Table 1 summarizes characteristics of the children and mothers in the estimation sample, additional information and summary statistics are given in table 8 in the appendix. The average child is almost eight years old and with a probability of 15% in full-day care whereas most children (58%) are cared of for only half of the day. 52% of all mothers participate in the labormarket, most of them in parttime work (39%). Those who work do so for on average 24 hours per week. A majority of 61% owns a vocational, 21% a tertiary degree. Descriptive statistics further indicate a positive correlation between maternal labor supply and full-day childcare. The mothers of children who are in care for the whole day work on average 6 hours per week longer (28.9 compared to 22.9). Similar, they exhibit higher labormarket participation rates (61% compared to 48%).

4.2 Data on All-Day Schools

Statistics on all-day schools in Germany as a whole are only available at the level of federal states. By using them one would impose the assumption that potentially very heterogeneous regions in a state are similar

¹² This is not a peculiarity of the sample used in this paper but rather a specific of the SOEP dataset (https://paneldata.org/variables/185145, February 1, 2017)

¹³ Computed from Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (2008), Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (2011) and Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (2016)
Figure 3: Share of all-day schools in all primary schools, state-level means of county-level data plotted against years by federal states. Illustration by author, based on self-collected data.

Figure 4: Share of students visiting an all-day school in all primary school students, state-level means of county-level data plotted against years by federal states. Illustration by author, based on self-collected data.
in terms of their supply with all-day schools. This assumption does not hold, for instance in 2014 there are counties in Schleswig-Holstein where 36% of all primary schools have an all-day offer as well as there are counties where this ratio is 78%.

The data is therefore collected at the county-level with the advantage of relatively good availability of school data as well as of other data to be included as control variables. Still, there is the potential problem that most primary schools in Germany are organized in school districts ("Schulsprengel"), a classification that is finer than counties. Parents are usually required to send their children to the school that is closest to their home, so the assumption that county-level data adequately depicts the all-day school supply in the counties’ school districts has to be made. This assumption rather circumvents some potential problems than creating them. First, the requirement to choose the nearest school does not apply in Schleswig-Holstein and North Rhine-Westphalia and is less strict in Brandenburg (together this affects about one third of all children in the sample). Second, states can allow exceptions from this requirement¹⁴. That could lead to a kind of “manipulation” as parents might want to have their children in a school that is nearer to their workplace than their home. A behavior that, if it is present, will arguably be more pronounced among working mothers, i.e. this paper’s group of interest. In this case, the county level as the broader category likely is a better representation of the all-day school supply mothers face.

The all-day school data is collected from the states’ statistical offices as well as from the ministries of education and, if necessary complemented with data from federal statistical databases¹⁵. From these data two variables are constructed. The share of primary schools with an all-day offer in all primary schools (abbreviated as “Share ADS”) and the share of students who participate in an all-day offer in all primary school students (“Share AD Students”); both at the county-level. As it solely depicts supply without including potentially endogenous decisions of families the share of all-day school is used as instrument in the main estimations while the students participating are used in a robustness check.

Table 2 lists the federal states for which data on all-day schools is available together with summary statistics. The dataset comprises information on the all-day school supply in most federal states in West Germany (only Lower Saxony and Bremen are missing) as well as

<table>
<thead>
<tr>
<th></th>
<th>Share of All-Day Schools</th>
<th>Share of Students in All-Day Schools</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Years</td>
<td>Mean</td>
</tr>
<tr>
<td>BW 07-14</td>
<td>0.15</td>
<td>0.1</td>
</tr>
<tr>
<td>BY 04-14</td>
<td>0.1</td>
<td>0.12</td>
</tr>
<tr>
<td>BE 02-14</td>
<td>0.84</td>
<td>0.25</td>
</tr>
<tr>
<td>BB 04-14</td>
<td>0.39</td>
<td>0.18</td>
</tr>
<tr>
<td>HH 02-14</td>
<td>0.29</td>
<td>0.33</td>
</tr>
<tr>
<td>HE 04-14</td>
<td>0.19</td>
<td>0.17</td>
</tr>
<tr>
<td>NW 04-14</td>
<td>0.72</td>
<td>0.27</td>
</tr>
<tr>
<td>RP 06-14</td>
<td>0.46</td>
<td>0.27</td>
</tr>
<tr>
<td>SH 02-14</td>
<td>0.27</td>
<td>0.21</td>
</tr>
<tr>
<td>SL 02-13</td>
<td>0.86</td>
<td>0.14</td>
</tr>
<tr>
<td>TH 02-14</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Overall</td>
<td>0.38</td>
<td>0.35</td>
</tr>
</tbody>
</table>

¹ State-level correlation between the share of all-day schools and the share of primary schools students who participate in an all-day offer.

Table 2: Summary of All-day School Data by Federal States.


¹⁵ Section 11.2 provides a description of the data.
in two East German states, Brandenburg and Thuringia. Even though views on childcare and female labor supply are traditionally different in the eastern part of Germany and Thuringia shows a full supply with 100% of primary schools being all-day schools these data is included. This is done because the IV-strategy chosen estimates a local effect of all-day schools on childcare and in turn labor supply such that – together with allowing for persistent regional heterogeneity as well as for a time trend – possible biases through effects of culture are minimized. Potentially remaining effect heterogeneity among the former inner-German border is explored in section 5.4.

Not all states started data collection in 2002, but for the largest part of the sample data for the years since 2004 are available. In Baden-Württemberg and Rhineland-Palatinate data were not collected before 2007, respectively 2006. As depicted in figure 3 the overall supply as well as how it developed over time is heterogeneous among federal states. While the overall mean share of all-day schools is 38%, the supply is some states as Berlin, North Rhine-Westphalia or, as noted, Thuringia is about twice as large (72 to 100%). In Baden-Württemberg, Bavaria and Hesse it only ranges between 10 and 19%. States as Hamburg, North Rhine-Westphalia or Rhineland-Palatinate show very steep increases (steepest in Hamburg where the supply increased from 3 to 97% between 2002 and 2014) whereas in others, for instance Bavaria or Saarland, the increase is moderate (the latter, however started at a higher level). The share of students who visit all-day schools, i.e. the take-up, is strongly correlated with the supply (the correlation overall is 0.81, coefficients at the state-level are listed in table 2) but generally lower with an overall mean of 19% (see also figure 4). In North Rhine-Westphalia or Rhineland-Palatinate the differences between strongly increasing supply and slower take-up is especially pronounced. Only in Hamburg, Brandenburg and Thuringia more than half of all primary school students visit an all-day school.

There are differences in the data compared to other work that also investigates all-day schools in Germany. Shure (2016) uses a dataset with similar information on schools but only for Bavaria, Hesse, Rhineland-Palatinate and Schleswig-Holstein in time range 2000 to 2012. Similar to Shure (2016) the dataset of Boll and Hoffmann (2017) is considerably smaller and both encode the availability of all-day schools not as precise shares but binary. Dehos and Paul (2016) use county-level school data which are drawn from information which schools received grants from the “Investment Program Future Education and Care”. This data source imposes some restrictions. As the program was in action between 2003 and 2009 only data for this period of time can be available. In addition, solely newly created all-day schools can be considered, not those already present. If counted nationwide 10.3% of German primary schools already had an all-day offer in 2002 (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland 2008). Likewise, all-day schools created after 2009 are omitted. While in 2009 (again measured nationwide) 41.7% of primary schools were all-day schools, this share increased to 53.3% in 2014. Moreover, Dehos and Paul (2016) have to make assumptions on how long it takes from when a school is given a grant until the all-day offer starts that may introduce further inaccuracy.

Compared to the existing ones the dataset used for this paper still has the shortcoming that not all federal states in all potentially interesting years are included. However, it is a substantial improvement in terms of comprehensiveness and precision. As it also uses the most recent available data the results of the analysis are likely to be most suitable to the current situation if implications for policy are to be drawn.

4.3 Other County-Level Data

The data on households and all-day schools are complemented with county characteristics. These are per capita gross domestic product (GDP p.c.), unemployment rates and population density. They together depict the socioeconomic characteristics of a county as well as if it is in an urban or rural area such that they provide an impression of employment possibilities for mothers.
All these data are collected from federal statistical databases.

More information on the county-level data together with summary statistics is provided in section 11.3 of the appendix.

5 Results

5.1 Descriptive Results

Before turning to the estimations this section illustrates the reduced form and first stage relations graphically. In all graphs the x-axis shows the share of all-day schools in a county, split in groups of 5 percentage points (pp) over which means of the dependent variable are calculated. A red line indicates a simple, bivariate fit.

The reduced form graphs of the relationship of all-day schools and labor supply (figures 5 and 6) are ambiguous. For both the intensive and the extensive margin the plots show several fractures with higher means of working hours and the labor market participation indicator respectively around 50% all-day schools and lower means around 80%. The overall trends, however, diverge. While the trend for working hours is positive (the slope of the linear fit amounts to around 3.5 hours), it is negative for labor market participation (the slope indicates a decrease of around 6 pp).

The graphical representation of the first stage (figure 7), i.e. how all-day schools affect if children are in full-day care, is, in comparison, much clearer. Apart from some positive outliers at around 40 and 50% all-day school supply the plot indicates a positive relation. The linear fit estimates this relation to be sizable at around 20 pp.

None of the graphs, however, show causal effects, they rather plot correlations as heterogeneity across households, counties or federal states as well as time trends are not taken into account.

5.2 Estimation-Results for the Full Sample

Table 3 lists the results for the estimation on the full sample. The dependent variables are weekly working hours for the intensive margin (columns I.1 to I.3) and an indicator that equals one if a mother works full- or parttime in the estimates for the extensive margin (columns E.1 to E.3). The models are set up according to the estimation equations [8] to [10] presented in section 3.3 such that the respective first columns estimate the simplest and the third ones the full model. Level of observation is always the child. As the intensive margin estimations are restricted to children whose mothers work the sample size is smaller compared to the extensive margin.

The first rows show the results of OLS models which are estimated as comparison. They indicate that working mothers who have their children in full-day care work about 3 to 4 hours per week longer while at the extensive margin the probability of being employed increases by 9 to 13 pp. Even though they are significant at the 1%-level throughout the results of OLS estimations cannot be assumed to be causal effects as discussed in section 3. Nevertheless, their size is of interest. The extensive margin estimates range around 10 pp and those for hours – assuming that full-time work amounts to 35 to 40 weekly hours – indicate increases in a similar range. As working mothers are more likely to have their children in care the OLS results include endogeneity that would result in coefficients that are biased upwards such that the estimated increases can considered to be small. Thus, the overall link between full-day childcare and maternal employment seems to be relatively weak.

The first stage estimates are crucial for the identification strategy. For the intensive margin the F-statistics range between 24 and 36, for the extensive margin between 22 and 60 where the larger maximum value is likely the result of a larger sample size. As they exceed both 10 as the commonly used rule of thumb as well as the tighter requirement of 16.38 proposed by Stock and Yogo (2005) the share of all-day schools in a county is a relevant instrument. The coefficients are highly significant throughout, their size is mostly constant at around
Figure 5: Graphical representation of the reduced form relation between all-day schools and weekly working hours at the county-level. The red line depicts a bivariate linear fit. Illustration by author, based on the estimation sample derived from the SOEP-data.

Figure 6: Graphical representation of the reduced form relation between all-day schools and labor market participation at the county-level. The red line depicts a bivariate linear fit. Illustration by author, based on the estimation sample derived from the SOEP-data.
<table>
<thead>
<tr>
<th></th>
<th>Intensive Margin</th>
<th>Extensive Margin</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Weekly Working Hours</td>
<td>Labormarket Participation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(I.1)</td>
<td>(I.2)</td>
<td>(I.3)</td>
</tr>
<tr>
<td><strong>OLS</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full-day Care</td>
<td>4.40***</td>
<td>3.47***</td>
<td>3.07***</td>
</tr>
<tr>
<td></td>
<td>(0.38)</td>
<td>(0.37)</td>
<td>(0.36)</td>
</tr>
<tr>
<td>R²</td>
<td>0.09</td>
<td>0.15</td>
<td>0.22</td>
</tr>
<tr>
<td><strong>First Stage</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Share ADS</td>
<td>0.21***</td>
<td>0.21***</td>
<td>0.20***</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.04)</td>
</tr>
<tr>
<td>R²</td>
<td>0.08</td>
<td>0.11</td>
<td>0.16</td>
</tr>
<tr>
<td>Robust F</td>
<td>34.26</td>
<td>35.52</td>
<td>23.87</td>
</tr>
<tr>
<td><strong>Reduced Form</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Share ADS</td>
<td>3.28**</td>
<td>2.57**</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>(1.32)</td>
<td>(1.24)</td>
<td>(1.6)</td>
</tr>
<tr>
<td>R²</td>
<td>0.073</td>
<td>0.14</td>
<td>0.22</td>
</tr>
<tr>
<td><strong>Two-Stage-Least-Squares IV</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full-day Care</td>
<td>15.73***</td>
<td>12.39**</td>
<td>4.94</td>
</tr>
<tr>
<td></td>
<td>(6.00)</td>
<td>(5.75)</td>
<td>(7.62)</td>
</tr>
<tr>
<td>R²</td>
<td>-</td>
<td>0.08</td>
<td>0.22</td>
</tr>
<tr>
<td>N</td>
<td>15298</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Control Variables**

- Person-level: no yes yes no yes yes
- County-level: no no yes no no yes

**Fixed Effects**

- Year: yes yes yes yes yes yes
- Federal State: yes yes no yes yes no
- County: no no yes no no yes

***/**/*** indicate significance at the 10%/5%/1%-level. Heteroskedasticity-robust standard errors are given in parentheses, in the OLS models clustered at the household-level, in the other models clustered at the county-level. Level of observation is the child. Share ADS measures the share of primary schools with an all-day offer in all primary schools in a county. Full-day care is a binary indicator if a child is in care for the whole day. Labormarket Participation is a binary indicator if a women works full- or parttime. Data sources are listed in section 11.

Table 3: Instrumental variable estimations of the effect of full-day childcare (instrumented by the share of all-day schools in a county, Share ADS) on maternal labor supply. Labor supply is measure at the intensive (weekly working hours) and extensive margin (indicator for full- or parttime work). Results are given for the full, unrestricted sample. Estimation equations are set up corresponding to equations [8] to [10].
Figure 7: Graphical representation of the first stage relation between all-day schools and children being in full-day childcare at the county level. The red line depicts a bivariate linear fit. Illustration by author, based on the estimation sample derived from the SOEP-data.

0.2 (with 0.17 in column E.3 still being in a similar range) indicating that for a hypothetical jump in the share of all-day schools from 0 to 1 the likelihood that a child is in full-day childcare increases by 20 pp. At the mean share of all-day schools (38%) this relates to an increase of 7.6 pp. This effect is not large, however, given that the requirements for all-day schools as well as the additional hours of care actually supplied (see section 2) are not necessarily sufficient to provide full-day childcare for the whole week an even smaller effect could had been expected.

While, so far, the results where mostly similar for the intensive and extensive margin and across specifications that changes in the reduced form. Estimates of the effect of all-day schools on labormarket participation are constantly negative, close to zero and, with standard errors almost of the size of the coefficients themselves, not significant. For the intensive margin the specification matters. As long as the county-level is not taken into account (columns I.1 and I.2) a statistically significant impact of all-day schools on maternal work can be found amounting to 2.6 to 3.3 hours per week for an (again hypothetical) increase in Share ADS from 0 to 1. If one, however, accounts for the fact that counties are generally different (by using fixed effects) and have different characteristics (by using covariates) no effect can be found any more. This is mostly the case due to a decrease in coefficient size while the standard errors change less. Overall, no convincing evidence for a robust ITT can be found.

The results of the full instrumental variable estimations show a pattern consistent with those of the reduce form. No causal effect of full-day childcare on labormarket participation of mothers that is statistically significant can be found. At the intensive margin size and significance depend on the specification. If for state and year effects as well as for household characteristics is controlled the LATEs range between about 15.7 and 12.4 additional weekly hours for mothers whose children are in care due to the existence of all-day schools, significant at 1%, respectively 5%-levels.

If, as in column (I.3), control variables at the county-level are added to the model the effect-size strongly decreases to about 4.9, such that, together with the standard error now being larger than the coefficient itself, no significant effect can be found any more.
The strong effect when moving from column (I.2) to (I.3) is, both in reduced form and IV estimations, mostly driven by a decrease in coefficient size rather than by increasing standard errors. Accounting for the county-level, therefore, rather removes variation from the model than introduces noise. Table 11 in the appendix lists the full estimation results, i.e. including covariates, for columns (I.3) and (E.3). As none of the county-level control variables has a significant impact those effects picked up by the county fixed effects seem to be more relevant. Thus, the main source of heterogeneity among counties are not economic characteristics as GDP or unemployment but time-invariant factors such as attitudes of the population and political factors (in turn shaped by attitudes). Since these factors potentially impacted where all-day schools were established, as also noted by Gambaro, Marcus, and Peter (2016), controlling for them is necessary to ensure that the independence assumption holds conditionally. Thus, the full specification in columns (I.3) and (E.3) is considered the most credible one. It, further, exhibits, in all models the highest R² values, i.e. it explains most of the models’ linear variation.

Even though, time constant differences at the county-level have an important effect in reduced form and IV estimates, their impact is less pronounced in OLS and first stage models. Therefore, it can additionally be concluded that parents in different counties do not show fundamental differences when they decide if they send their children to full-day care or in how they translate the fact that their children are in care into labor supply. This finding is still consistent with the conclusion of differences in, e.g., policymaking as parents are neither the only group that might affect political decision nor the one that is necessarily organized best.

Overall, the estimations on the full sample do not find robust significant effects of all-day schools and in turn full-day childcare on maternal labor supply. This finding is consistent over all specifications for labormarket participation while at the intensive margin county-level factors play a role. Before drawing conclusion the next section explores the role of the all-day school availability in more depth.

5.3 Estimation-Results for States Where All-Day School Supply is High

The findings in the full sample estimations show the presence of impacts on the county-level. Since in terms of control variables the chosen approach of adding county characteristics as well as county fixed effects already allows for very much flexibility now the size of the supply with all-day schools itself is to be investigated. It is not only taken into account by the according variable, but additionally the sample is split in those federal states where on average (over counties and years) less than 20% of primary schools have an all-day offer and those with more than 20% all-day schools and the estimation is carried out twice. The sample with low all-day school supply consists of Baden-Württemberg, Bavaria and Hesse, that with high supply accordingly of all others (see also table 2). Using 20% as cutoff to split the sample ensures that the resulting groups are of similar size and thus comparable. The state is chosen as the level to split the sample to obtain groups which are consistent, such that a county is either in one subsample for the whole period of observation or in the other. Using the county-level supply as reference point would have the result that a number of counties switches from one to the other subsample as more all-day schools are created over time. Thus, eventually a time trend would be, in addition to all-day school supply, an important differentiating factor. The results of these estimations are listed in table 4. In all cases the full specification with all covariates as well as time and county fixed effects according to equation [10] is estimated.

All OLS estimates yield positive and significant coefficients. Their size is smaller for states with low supply with all-day schools and larger for the other subsample. They are, nevertheless, in a similar range, also if compared with the full-sample estimates from table 3. It therefore can be concluded that mothers in the low all-day school supply sample while showing an overall less strong reaction to their children being in full-day care are not fundamentally different[^16].

[^16]: This conclusion can be drawn despite the likely presence of an endogeneity bias as it would either affect both groups in a similar way.
<table>
<thead>
<tr>
<th></th>
<th>Intensive Margin</th>
<th>Extensive Margin</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Weekly Working Hours</td>
<td>Labormarket Participation</td>
</tr>
<tr>
<td>Share ADS &lt; 20%</td>
<td>Share ADS &gt; 20%</td>
<td>Share ADS &lt; 20%</td>
</tr>
<tr>
<td>(I.1)</td>
<td>(I.2)</td>
<td>(E.1)</td>
</tr>
<tr>
<td><strong>OLS</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full-day Care</td>
<td>2.56***</td>
<td>3.33***</td>
</tr>
<tr>
<td></td>
<td>(0.66)</td>
<td>(0.42)</td>
</tr>
<tr>
<td><strong>R²</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First Stage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Share ADS</td>
<td>0.06</td>
<td>0.27***</td>
</tr>
<tr>
<td></td>
<td>(0.1)</td>
<td>(0.06)</td>
</tr>
<tr>
<td><strong>R²</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Robust F</td>
<td>0.1</td>
<td>0.15</td>
</tr>
<tr>
<td>Reduced Form</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Share ADS</td>
<td>-8.54***</td>
<td>3.77*</td>
</tr>
<tr>
<td></td>
<td>(3.36)</td>
<td>(1.96)</td>
</tr>
<tr>
<td><strong>R²</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Two-Stage-Least-Squares IV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full-day Care</td>
<td>-149.12</td>
<td>13.72**</td>
</tr>
<tr>
<td></td>
<td>(253.9)</td>
<td>(6.98)</td>
</tr>
<tr>
<td><strong>R²</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>7385</td>
<td>7913</td>
</tr>
</tbody>
</table>

Control Variables: Person/household- and county-level
Fixed Effects: Year and county

*/**/*** indicate significance at the 10%/5%/1%-level. Heteroskedasticity-robust standard errors are given in parentheses; in the OLS models clustered at the household-level, in the other models clustered at the county-level. Level of observation is the child. Share ADS measures the share of primary schools with an all-day offer in all primary schools in a county. Full-day care is a binary indicator if a child is in care for the whole day. Labormarket Participation is a binary indicator if a women works full- or parttime. Data sources are listed in section 11.

Table 4: Instrumental variable estimations of the effect of full-day childcare (instrumented by the share of all-day schools in a county, Share ADS) on maternal labor supply. Labor supply is measure at the intensive (weekly working hours) and extensive margin (indicator for full- or parttime work). Estimations are carried out separately for federal states with on average less (Baden-Württemberg, Bavaria, Hesse) and more (all others) than 20% all-day schools. In all cases the full specification according to equation [10] is estimated.
The differences between the samples become more pronounced in the first stage. For both the intensive as well as the extensive margin in the samples with larger all-day school supply (columns I.2 and E.2) the coefficients on Share ADS are larger than in the full sample (0.27 and 0.23 compared to 0.2 and 0.17). In either case they are highly significant. In the low supply sample (columns I.1 and E.1) these coefficients are considerably smaller (0.06 and 0.04) and not significant. Correspondingly, all-day schools are only relevant instruments if their supply is higher (with F-statistics of at least 23 compared 0.4). The coefficient in column (I.2) indicates an increase of the probability that a child is in full-day care by 27 pp if the share of all-day schools increases from 0 to 1. Compared to the first stage estimates in the full sample this increase seems moderate, however, now the mean of Share ADS is not 38% but 65%. Thus, at the mean all-day school supply the first stage estimates an increase in the probability of full-day care by 17.6 pp, more than twice the size of the effect in the full sample (which amounted to 7.6 pp).

Both reduced form results for the low supply sample show effects that are negative and highly significant. Additionally, these effects are large compared to the corresponding results for the sample with higher all-day school supply. A finding that neither matches expectations nor potential theoretical explanations of how all-day schools affect labor supply. While no extensive explanation can be given there are potential reasons that can be ruled out. It does most likely not indicate that the relation of full-day childcare and maternal labor supply is fundamentally different between the two subsamples as as shown by the OLS results being in comparable ranges. Based on the considerations in section 3.2 also a violation of the monotonicity assumption in states with low all-day school supply does not seem plausible. Further, as for the economic situation of counties is controlled, also the potential explanation that in the low supply sample all-day school were especially (and disproportionately compared to the high supply sample) or - if preferences of parents in the samples were different such that those in states with higher all-day school supply have a higher preference for work - would rather increase the difference between the estimates targeted to economically weak counties is hard to maintain.

Together with the negative reduced forms the instrumental variable estimations for the low supply sample do not find significant effects, the coefficients are even negative. The extensive margin estimate on the high supply sample is positive but not significant.

If, however, the intensive margin in states with at least 20% all-day schools is considered, the highly significant first stage along with a slightly significant reduced form lead to an effect of 13.72 additional working hours per week if a child is in full-day care, significant at the 5%-level. This is a local estimate for those mothers who have their children in care due to they existence of all-day schools. The more general ITT that is estimated by the reduced form is smaller (3.77 additional weekly hours) and only slightly significant. Compared to OLS results and ITT the LATE estimated via IV is about four times larger. This highlights that, in comparison to ITT, LATE not considers all mothers who could utilize all-day schools but rather solely those who in fact do. It further indicates that those mothers who send their children to all-day schools show a stronger reaction to their children being in full-day care than working mothers in general. A potential explanation might be that mothers who prefer all-day schools over other means of childcare have a higher preference for work. This work preference could, in turn, be related to a higher concern about their children’s education and future chances in the labor market what would explain why they use all-day schools.

There are further conclusions that can be drawn from the results obtained so far. Even though, all-day schools do contribute to children being in full-day care and can increase hours worked if their supply is sufficiently high, they are in neither case capable of making mothers change their labor market participation status. Likely, this is because the additional hours of care provided do not in all cases relate to what one might expect of “full-time” care (see chapter 2) such that those mothers who decide to (re-)enter the labor market when their children enter primary school utilize other means of childcare. Moreover, finding no effects on the extensive margin
is consistent with the explanation relying on work preferences given above, since, if all-day schools are mostly used by mothers with higher work preference, they are more likely to already participate in the labor market when school starts.

The finding that effects on the intensive margin can only be observed if all-day school supply is higher than on average 20% can have several reasons. It, first, could simply be the consequence of the relatively small sample size in the SOEP. In 2014 over 2.7 million students in Germany visited a primary school (Statistische Ämter des Bundes und der Länder 2016a), about one third of them an all-day school. The full sample used for this paper consists of about 15,300 children (considering the intensive margin estimates), i.e. a 0.57% subsample of the overall population that is not fully representative. For instance, there are 31% of mothers with a migration background in the sample¹⁷ while the overall share in 2014 only was 20.3%¹⁸. Migrants arguably are different in terms of labor supply so they are an example of how estimations might be biased downwards. Second, a more general sample size effect would be that due to few observations it is harder to catch those who utilize all-day schools for childcare and increase their labor supply if the share of all-day schools is below a certain threshold, in the present case 20%. This hypothesis of a sample size effect – regardless of its particular nature – is, however, is to some extend contradicted by the results of Dehos and Paul (2016) who use a much larger sample from the German Microcensus (with more than 140,000 observations) to compute the reduced form and still do not find a significant effect. That supports the third possible explanation that finding a significant effect on hours above a share of all-day school of 20% is not a problem of measuring but rather directly linked to supply. It might be that an all-day school share of more than 20% indicates that expansion already started in previous years giving parents time to, first, know about the possibility and, second, make according decisions. Or, without making assumptions about the timing of the expansion, parents might need a large enough supply to be sufficiently sure to get a place in an all-day school and then make decisions. Eventually, it has to be noted that solely considering federal states with more than 20% all-day school supply removes Baden-Württemberg, Bavaria and Hesse from the sample. These are states that are traditionally more conservative. Even though, the OLS estimates do not find strong evidence for fundamental differences between parents in the samples an impact of preferences at the state-level cannot fully be ruled out. It even gains some support from differential reduced form estimates. As Baden-Württemberg, Bavaria and Hesse are the richest German states¹⁹ it is also possible that effects at the state-level are not associated with preferences but rather with characteristics of the population such that a potentially richer population in richer states exhibits different behavior.

5.4 Complementary Results

In addition to the main results given above this section explores the heterogeneity of the effects along two dimensions.

In contrast to Shure (2016) and Dehos and Paul (2016) this paper does not only concentrate on West Germany but also uses data from two states in the former German Democratic Republic, Brandenburg and Thuringia. Views on childcare and female labor are traditionally different in East Germany, a specific that can be found in this paper’s estimation sample, too. While in the western federal states on average 13% of 5 to 11 year olds are in full-day care this share is 35% in the East. Likewise, 60% of women in the two eastern states work, compared to 51%; those who work do so for on average 33 weekly hours, that is about 10 hours more than in West Germany.

Therefore, it is, first, examined if and how the results differ between the eastern and western states in the sample. This is done by splitting the sample into East and West

---

¹⁷ As Wagner, Frick, and Schupp (2007) note, too, the SOEP oversamples migrants
¹⁸ Own computation using https://www.destatis.de/DE/ZahlenFakten/GesellschaftStaat/Bevoelkerung/MigrationIntegration/Migrationshintergrund/Tabellen/MigrationshintergrundGeschlecht.html, February 2, 2017
¹⁹ They are currently the only donor states in the “Federal Financial Equalisation System” (Länderfinanzausgleich); Bundesministerium der Finanzen (2016)
and repeating the estimation of the full specification for the intensive margin. The results, listed in columns (1) (East) and (2) (West) of table 5, are to be compared with column (I.3) of table 3. Contrasting the overall higher utilization of childcare and the fact that there is unconstrained supply with all-day schools in Thuringia in East Germany all-day schools are not a relevant instrument for childcare (F-statistic of 0.39) while they are in the West (F-statistic of around 27). The first-stage coefficient for the East is larger, 0.24 compared to 0.19, but is exceeded by its standard error whereas the one for the western states is close the full sample estimation and highly significant. Matching the results from the full sample the IV estimation does not find significant effects in both subsamples. While the estimate for the West is still positive and in the range of that from the full sample, the IV coefficient for the East even indicates a negative relation of full-day care and maternal working hours. These results have to be handled with care as the subsample for West Germany is about ten times larger, however they give hints that the effect of all-day schools on children being in care is more robust in the West and, in turn, the effect on labor supply is stronger (if only signs, not significance of the coefficients is considered). Arguably, this is related to a generally better supply with childcare in East Germany, further tests with a more comprehensive data base are nevertheless necessary to explore why, despite their availability, all-day schools are not more often utilized.

Second, heterogeneity along education of mothers is analyzed. According to the descriptive results of Marcus, Nemiță, and Spieß (2013) with their expansion all-day schools were increasingly visited by children from lower income households while overall, i.e. if no time trend is considered, children of mothers with a university degree are more likely to go there. To further investigate this issue columns (3) and (4) of table 5 split the sample into mothers who have completed university-level education and those who have a vocational or no degree. Using measures of education to split the sample implicitly includes other factors as income or socio-economic status in general since these are usually positively re-
lated to education. Contrary to the first to columns where the full sample was split, here only the restricted sample of states with at least 20% all-day schools (i.e. Baden-Württemberg, Bavaria and Hesse are dropped) is used as it generated the so far only significant result. Again, the intensive margin is considered such that the estimates are to be compared with column (I.2) in table 4.

For children whose mothers own a tertiary degree the first stage estimates do not find an impact of all-day schools on full-day care, the coefficient is insignificant and the instrument not relevant. Consequently, no significant effect can be found in the IV estimates. In the less well educated subsample, in contrast, all-day schools have a significant impact on children being in full-day care, with a coefficient size of 0.29 (which translates to an increase in the probability of full-day care by 29 pp for an increase in the all-day school supply from 0 to 1 or 18.9 pp at the mean all-day school supply). IV estimates for this subgroup find an effect of around 20 additional weekly working hours for mothers whose children are in full-day care due to all-day schools, significant at the 5%-level.

These results match to some extend differences between the two samples. If mothers own a university degree their children are more likely to be in full-day care (utilization rate of 18% compared to 15%). Better educated mothers additionally show an on average higher labor supply. Among those who work the ones with a university degree work more hours (28 compared to 23) and are more often in fulltime work (in the highly educated subsample the mothers of 30% of the children work full-time, in the other subsample this share 21%). Such education specific differences in labor supply are not specific to the estimation sample but can also be found in larger datasets as the Microcensus (Bundesministerium für Familie, Senioren, Frauen und Jugend 2014, p. 33). This indicates that the initial potential to utilize all-day schools as well as to increase labor supply is restricted among well educated mothers. However, while the estimation results might be explained partly by different initial situations of education groups the are likely to primarily reflect a general difference in how all-day schools are utilized. As they have – on average – lower earnings (in the estimation sample their hourly wage is with 11.60€ almost 8€ smaller than that of mothers with university degree) less educated mothers may prefer all-day schools as they offer, besides care, promotion of the children (as homework assistance) for which the parents in other cases had to pay. Overall, this result suggests that all-day schools can be an effective instrument to reduce the employment gap between differentially educated mothers.

6 Robustness Checks

This section discusses potential threats to the estimation. Three checks for the main significant finding, that all-day schools have an impact on the intensive margin of labor supply if their supply is sufficiently large, are presented. Further approaches address the relevance of all-day schools as an instrument as well as a possible endogeneity problem.

Columns (1) to (3) of table 6 repeat the intensive margin estimation on the sample restricted to federal states with on average 20% or more all-day schools. They are therefore to be compared to column (I.2) of table 4.

As section 4.1 discusses, identification of children in primary school relies their age. Specifically, those between 5 and 11 are assumed to be in primary school. This assumption is tested by narrowing the age range of children in the estimation sample to those who are between 6 and 10 years old. With the smallest and biggest value of the age distribution dropped the remaining children can easier be assumed to be in primary school. The results of this test are listed in column (1) of table 6. With a size of 0.26 (compared to 0.27 in the estimation with 5 to 11 year olds) and significance still at the 1%-level first stage results are virtually unchanged. The coefficient on the full-day care indicator in the IV estimation increases slightly to 14.0 (compared to 13.72) while its standard error increases relatively stronger resulting in an estimate only significant at the 10%-level. As this test reduces the number of observations by about one third (5641 compared to 7913) its findings are arguably to a
<table>
<thead>
<tr>
<th>Dependent Variable: Weekly Working Hours</th>
<th>(1) Age Range 6–10</th>
<th>(2) Control for Other Childcare Facilities</th>
<th>(3) Without Education and Childcare Sector</th>
<th>(4) Half-Day Care as Explanatory Variable</th>
<th>(5) Students in All-Day Schools as Instrument</th>
</tr>
</thead>
</table>

**First Stage**

<table>
<thead>
<tr>
<th>Share ADS</th>
<th>0.26***</th>
<th>0.32***</th>
<th>0.29***</th>
<th>-0.13***</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0.06)</td>
<td>(0.08)</td>
<td>(0.06)</td>
<td>(0.04)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Share AD Students</th>
<th>0.09</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(0.08)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>R²</th>
<th>0.18</th>
<th>0.16</th>
<th>0.15</th>
<th>0.36</th>
<th>0.17</th>
</tr>
</thead>
<tbody>
<tr>
<td>Robust F</td>
<td>16.23</td>
<td>20.09</td>
<td>26.39</td>
<td>8.53</td>
<td>1.06</td>
</tr>
</tbody>
</table>

**Two-Stage-Least-Squares IV**

<table>
<thead>
<tr>
<th>Full-day Care</th>
<th>14.0*</th>
<th>19.26*</th>
<th>12.54*</th>
<th>-7.15</th>
</tr>
</thead>
<tbody>
<tr>
<td>(8.07)</td>
<td>(10.0)</td>
<td>(6.82)</td>
<td>(37.94)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Half-day Care</th>
<th>-7.77</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(12.31)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>R²</th>
<th>0.13</th>
<th>-</th>
<th>0.15</th>
<th>0.2</th>
<th>0.14</th>
</tr>
</thead>
</table>

| N  | 5641 | 6401 | 6927 | 15298 | 12303 |

**Control Variables:** Person/household- and county-level

**Fixed Effects:** Year and county

* / ** / *** indicate significance at the 10%/5%/1%-level. Heteroskedasticity-robust standard errors, clustered at the county-level, are given in parentheses. Level of observation it the child. Share ADS measures the share of primary schools with an all-day offer in all primary schools in a county. Share AD Students measures the share of primary school students participating in an all-day offer. Full-day care is a binary indicator if a child is in care for the whole day. Half-day Care is its counterpart measuring if a child is in care for half of the day. Data sources are listed in section 11.

**Table 6:** Instrumental variable estimations that check the robustness of the previous results. Dependent variable is always labor supply at the intensive margin (working hours per week), endogenous explanatory variable is full-day or half-day (column 3) childcare, the instrument is share of primary schools with an all-day offer in a county (Share ADS) or the share of students in all-day schools (Share AD Students, column 4). In all cases the full specification according to equation [10] is estimated.
large extend the result of a smaller sample size rather than they would indicate a systematic bias that was introduced by including 5 and 11 year olds.

In column (2) the fact that estimations so far did not control for childcare facilities other than all-day schools (Horte) is addressed. Section 3.2 discusses the weaknesses of the data on other childcare facilities and that they are not crowded out by all-day schools. Still, it might be possible that using county fixed effects and allowing for a time trend is not sufficient to depict the availability of Horte in a county adequately. Therefore, the estimation contains a control variable measuring the total number of childcare places not supplied by all-day schools in a county, normalized by the number of children below the age of ten. Places are chosen as the are, in contrast to facilities, easier to adjust to demand, thus arguably the most precise measure. Even though, the coefficient increases, indicating 19.26 additional weekly hours of work if a child is in full-day care, it looses some accuracy, being significant at a 10%-level. The overall trend of the estimate, however, does not change. Notably, the first stage coefficient increases to 0.32, the impact of all-day schools on full-day childcare is, thus, robust. Neither in the first stage nor in the IV estimation the number of childcare places shows a significant impact (for economy of space not listed in the table), a circumstance that supports the initial decision to exclude this variable in the main estimations.

The test in column (3) examines if there is, besides the indirect effect through the provision of childcare, some different direct impact of all-day schools on maternal labor supply, thus if the exclusion restriction can be assumed to hold. It is likely that, since they get expanded, all-day schools have an increasing demand for caregivers. If this demand is satisfied by mothers who take up a job in an all-day school, and in turn increase their labor supply, that would establish a new channel of causality. The estimation in column (3) of table 6 is therefore computed without the children whose mothers work in the childcare or education sector (an approach proposed by Shure 2016) such that the additional channel of impact, if present, is removed. The IV coefficient indicates an effect of 12.54 additional working hours if a child is in full-day care, a decrease of about 1.2 hours compared to the unrestricted specification. Even though, the standard error decreased slightly, this results is now significant only at the 10%-level. That indicates that there may indeed be some mothers working as caregivers in all-day schools who increased their labor supply. However, the overall range of the estimates stays similar, so those mothers cannot be considered a main driving factor of the results obtained in table 4. As opening hours and working hours match almost perfectly one would expect that especially those mothers who potentially work in all-day schools utilize them for childcare more often. No evidence for this expectation can be found. The first stage coefficient on the impact of all-day schools on full-day childcare is now larger than in the sample which includes mothers who work in education and childcare (0.29 compared to 0.27) suggesting that those who not work there are more likely to use all-day schools. Overall, this test does not indicate that increased employment of mothers in all-day schools severely biases the estimates.

Even though they proof to be a relevant instrument the impact of all-day schools is limited. The result in section 5.2 indicates an 7.6 pp increase in the likelihood of a child being in full-day care at the mean share of all-day schools, if the estimates of section 5.3 are considered this effect raises to 17.6 pp. The amount of care all-day schools provide is also limited as discussed in section 2 which is the likely reason that no effects on the extensive margin can be found. These two points together might raise the concern that the main channel of impact through which all-day schools have an effect is, despite their name, not full-day but in fact half-day childcare. For instance, that could mean, that only in some states where they offer more hours of care (as found by Klemm and Zorn 2016) they are utilized for full-day care while in others they had their impact through the channel of half-day care. The result was, as only one part of the effect was considered, an underestimation of how all-day schools contribute to labor supply. This concern is assessed in column (4) of table 6 by repeating the full sample estimation from table 3, column (I.3), but now an indicator if a child is in half-day care is used as en-
ogenous explanatory variable. Of primary interest in this estimation is the result of the first stage. With a F-statistic of 8.53 the share of all-day schools is a weak instrument for half-day childcare, both in absolute terms (as the F-statistic is smaller than 10) as well as in comparison to estimates where full-day care is to be explained. Even if relevance would be satisfied, the effect of all-day schools on half-day care is estimated to be negative and significant at the 1%-level, with a coefficient of -0.13 about half as large (in absolute values) as if full-day care is considered. This suggests that in counties with more all-day schools children are more often in full-day than in half-day care. Accordingly, the instrumental variable estimation finds a negative effect of half-day care on working hours, which is, however, not significant. Based on the results of this test the hypothesis that all-day schools might have impacts different than through full-day childcare can be rejected.

Section 5.3 provides evidence for the hypothesis that a sufficiently large supply with all-day schools is crucial for their contribution to maternal labor supply. Still, there is the possibility that supply is not the causal factor but rather itself driven by something different. The underlying mechanism might be that all-day schools were especially built where parents indicated that they wanted to have them. That could be by campaigns of schools’ parents’ associations, by forming citizens’ initiatives or simply by voting for parties which are in favor of all-day schools in county- or state-level elections. Parents who are strongly in favor of all-day schools might additionally exhibit different preferences concerning their childcare and work decisions – a presumable reason why they support all-day schools – for which not necessarily can be controlled with the included variables. In such a situation the independence assumption of random assignment of the instrument would, even conditionally, not hold any more. An omitted variable bias in both first stage and reduced form would be the consequence. Even though it is not a rigorous test the following approach provides indication if unobserved parental preferences are present. It builds on the assumption that parents’ preferences can be proxied for by the share of primary schools students who participate in an all-day offer (Share AD Students), i.e. that a higher preference for all-day schools results in a higher utilization of them. This share of students who visit all-day schools is now used as an instrument for the indicator if a child is in full-day care. As listed in table 2 this variable is not available for all federal states and years which results in a smaller sample. If it was a relevant instrument and the corresponding IV estimates exhibited a significant effect it would indicate that in counties where mothers have a general preference in favor of all-day schools they utilize them more often and accordingly increase their labor supply. Column (5) of table 6 lists the results of this check. The estimation corresponds, except of the different choice of the instrument, to this in column (1.3) of table 3. No evidence for the described hypothesis can be found. The share of students who participate in all-day schools neither has a significant impact on a child being in full-day care nor it is a relevant instrument (F-statistic of 1.06). Consequently, no effect on working hours can be found. Therefore, no indication for parental preferences which impacted the supply with all-day schools can be found.

7 Discussion and Conclusion

This paper analyzes how all-day schools contribute to maternal labor supply in Germany. By using the county-level share of primary schools with an all-day offer as an instrument for full-day childcare a local average treatment effect is estimated, i.e. the results give an effect of full-day childcare on labor supply for the mothers of those children who are in care because of the presence of all-day schools.

The results indicate that the size of all-day school supply is a crucial factor for their effect. This paper differentiates between an estimation on the full sample, i.e. containing all federal states for which data are available, and one where the sample is split in states with low (share of all-day schools of on average smaller than 20%) and high (more than 20%) supply with all-day schools. The first stage results show that all-day schools have a relevant impact on childcare utilization. At the mean share of all-day schools this impact amounts to increases
in the probability that a child is in full-day care of 7.6 pp in the full sample and 17.6 pp in the sample with high supply. The impact in states with high supply is similar to the mean share of students who visit all-day schools of 19%. IV estimations on the full sample neither find an effect on the intensive nor on the extensive margin. If, however, the sample with states with high supply is considered the picture is a different one. While still no effect on labor market participation can be found, estimates for the intensive margin indicate a LATE of additional 13.7 weekly working hours for those mothers who have their children in full-day care. This result is significant at the 5%-level and proofs mostly robust in various specification checks. Further, an effect of otherwise unobserved preferences of mothers proxied for by the actual utilization rates of all-day schools can be ruled out. The ITT of the impact of all-day schools on the labor supply of all mother of primary school students estimated as the reduced form on the high supply sample is considerably smaller than the LATE (3.7 hours) and only slightly significant at the 10%-level.

The absence of a significant effect on the extensive margin confirms the results obtained by Dehos and Paul (2016), but is contrary to those of Gambaro, Marcus, and Peter (2016), Shure (2016) or the ex-ante simulation of Beblo, Lauer, and Wrohlich (2005). It is, however, consistent with the design of the German all-day schools system where schools do not have to provide care on all weekdays for the full day to qualify as "all-day". Assuming that mothers who consider entering the labor market utilize other means of childcare would to some extent explain the differential findings of Gambaro, Marcus, and Peter (2016) as they do not differentiate between care in all-day schools and care in other facilities (Horte).

While the ITT estimates on the high supply sample are in a range similar to those of Boll and Hoffmann (2017) the corresponding LATE of 13.7 additional weekly hours is large compared to other work. Interestingly, it matches the results of Klemm and Zorn (2016) who estimate that all-day schools on average provide 13.7 hours of additional care per week compared to half-day schools. This finding promotes the hypothesis that if all-day schools are sufficiently available those mothers who utilize them translate all the additional time their children are cared of to additional working hours almost in a one-to-one relation. Before considering it confirmed such an arguably strong conclusion, however, requires attention of further research.

The general finding that all-day schools only have an effect if their supply is sufficiently high suggests that a further expansion is advisable to make sustainable improvements to maternal labor supply. As discussed, parents may need to face a sufficiently high supply to be sure to get a place for their child such that they can make according decisions. These paper’s findings, therefore, support the hypothesis that in the current situation they are constrained by too little supply. This conclusion is supported by Hollenbach-Biele and Zorn (2016) who summarize the results of a representative survey among about 4300 parents of children between 6 and 16. In this survey 32% of parents whose children do not visit an all-day school state insufficient availability as a reason and that they would wish for higher supply.

Assuming a further expansion of all-day schools such that the effects found in the high supply sample generalize to Germany as a whole allows to calculate the approximate fiscal effect. In this scenario the share of students visiting an all-day school increases by 17.6 pp and their mothers work 13.7 hours per week more. Given that there are about 2.7m primary school students in Germany around 477,000 in addition would visit an all-day school. On the level of the mothers this translates to an increase in the expected yearly income tax payments of about €3,145 per mother and an overall increase of about €1.08bn. In addition to this increase in tax revenue also costs have to be considered. Klemm and Zorn (2016) estimate the average yearly personnel cost of one all-day class in the completely integrated type to amount €21,500. As this type of all-day schools is the least frequent and the one with the strictest attendance requirements costs for the other types can assumed to be lower such that this number likely constitutes a upper boundary. With the assumption of 25 students per class the

---

20 Details on the calculation can be found in section 13 of the appendix.
additional costs would add up to around €410m. The overall result of this calculation is a fiscal surplus of around €670m. However, this does not consider maintenance costs for schools and a number of fiscal benefits. For instance, mothers’ wages are likely to increase in labor supply due to human capital effects of work experience, they contribute more to social security and pension systems and will have higher consumption expenditures. Further, potentially beneficial effects on students, as shown by Boll and Hoffmann (2017) or StEG-Konsortium (2010), which in the long run translate to human capital and fertility effects of childcare (Bauernschuster, Hener, and Rainer 2016) have to be taken into account. Eventually, as the mothers of almost half a million of students would be affected, there is the possibility of general equilibrium effects such that maternal labor supply does not necessarily meets a sufficient demand, crowds out other (potentially paternal) employment (as the results of Felfe, Lechner, and Thiemann (2013) suggest) or leads to a decrease in wages have to be considered. A more comprehensive analysis of the fiscal effects of investments in childcare is provided by Krebs and Scheffel (2016) who conclude overall large positive effects on GDP, employment, wages and national budgets.

Even though this paper can provide some insights it also points out where further research can improve. With the ongoing expansion of all-day schools the possibility that they crowd out other providers of childcare after school increases such that controlling for the availability of these other childcare facilities becomes more important. This paper is planned to be extended in several ways. The variation created through the different levels of all-day school supply across federal states (Klemm and Zorn 2016) will be exploited to provide an accurate depiction of circumstanced faced by parents when making childcare and labor supply decisions. Further, more rigorous tests of the independence-assumption will be provided allowing to analyze potential driving factors behind all-day school creation at the local level.

8 References


Contreras, Dante, Paulina Sepúlveda, and Soledad Cabrera (2010). "The effects of lengthening the school day on female labor supply: Evidence from a quasi-experiment in Chile".
Hollenbach-Biele, Nicole and Dirk Zorn (2016). *Wie Eltern den Ganztag sehen: Erwartungen, Erfahrungen,*


Appendix

9 Additional Figures on the Institutional Background

See figures 8 and 9.

10 Test of the Monotonicity Assumption Required for LATE

The regressions in table 7 provide a test if the increased supply of all-day schools crowded out other means of childcare provision. The related discussion can be found in section 3.2.

11 Additional Information on the Data

11.1 SOEP-Data

From the different datasets available through the SOEP, the “SOEP-long” datasets, version 31.1 (SOEP 2014), are used. They provide a preprocessed version of the data in a panel format that is easier accessible. From the different datasets
Figure 8: Mean weekly working hours of mothers in Germany. In contrast to figure 1 measured per mother and depending on the age of the youngest child.
Bundesministerium für Familie, Senioren, Frauen und Jugend (2014, p. 38)

Information on mothers, male partners (if present) and children were retrieved and merged using household and person as well as year identifiers. Retrieving data from the SOEP and initial preparation builds on work by Holger Stichnoth (Centre for European Economic Research, Mannheim). To keep the estimations consistent in terms of control variables hourly wages for those women who do not work were imputed. For this imputation a regression of the known log wages on a variety of personal characteristics (education, migration background, job characteristics, number of children, martial status as well as age and experience in fulltime and parttime work and years in unemployment and according squared terms) and time as well as state fixed-effects was computed. The final dataset of children complemented with information on their parents was merged county-level data using county identifiers (“Kreiskennziffern”) and year identifiers. All data preparation and estimation was done via SOEPremote, a remote access service that allows to use the SOEP including information on the county where a household is located which are otherwise unavailable due to privacy protection laws.

Weekly working hours contain overtime hours if they are paid or not compensated for with free time. The full- or parttime indicator does not include marginal employment. If not imputed hourly wage is calculated from gross labor earnings. Nonlabor income contains income from investments and renting and, for mothers in couples, the income of the partner.

Table 8 lists additional summary statistics, table 9 shows how the observations (at the child-level) are distributed over federal states and years.

11.2 All-Day School Data

The data on all-day schools and – if available – students in all-day schools are collected from the ministries of education in the federal states, from the states’ statistical offices or drawn from the publications from the Kultusministerkonferenz. In cases where only the numbers of all-day schools and/or participating students were supplied the overall numbers of primary schools and students was drawn from Statistische Ämter des Bundes und der Länder (2016a) to compute shares on the county-level.

Information on the quality of the data are only available for North Rhine-Westphalia. Here, it is possible that for one to at most three schools per year in the years 2005 until 2010 sporadic double counts happened. The statistical office rates this issue as minor.
### Figure 9: Additional care-time supplied by all-day schools. Days with all-day offer per week, duration of these days and overall additional time per week. Source: Klemm and Zorn (2016, p. 28)

<table>
<thead>
<tr>
<th>Bundesland</th>
<th>Ganztage Anzahl pro Woche</th>
<th>Umfang der Ganztage Zeitstunden pro Ganztag</th>
<th>Mehrzeit an Ganztagen Zeitstunden pro Woche</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baden-Württemberg</td>
<td>4</td>
<td>8</td>
<td>14.9</td>
</tr>
<tr>
<td>Bayern</td>
<td>4</td>
<td>8</td>
<td>13.8</td>
</tr>
<tr>
<td>Berlin</td>
<td>4</td>
<td>8</td>
<td>16.2</td>
</tr>
<tr>
<td>Brandenburg</td>
<td>5</td>
<td>8</td>
<td>14.0</td>
</tr>
<tr>
<td>Bremen</td>
<td>5</td>
<td>8</td>
<td>16.4</td>
</tr>
<tr>
<td>Hamburg</td>
<td>5</td>
<td>8</td>
<td>22.4</td>
</tr>
<tr>
<td>Hessen</td>
<td>4</td>
<td>8</td>
<td>10.0</td>
</tr>
<tr>
<td>Mecklenburg-Vorpommern</td>
<td>3</td>
<td>8</td>
<td>8.7</td>
</tr>
<tr>
<td>Niedersachsen</td>
<td>4</td>
<td>8</td>
<td>15.1</td>
</tr>
<tr>
<td>Nordrhein-Westfalen</td>
<td>4</td>
<td>8</td>
<td>14.2</td>
</tr>
<tr>
<td>Rheinland-Pfalz</td>
<td>4</td>
<td>8</td>
<td>8.5</td>
</tr>
<tr>
<td>Saarland</td>
<td>3</td>
<td>8</td>
<td>7.9</td>
</tr>
<tr>
<td>Sachsen</td>
<td>4</td>
<td>8</td>
<td>16.9</td>
</tr>
<tr>
<td>Sachsen-Anhalt</td>
<td>3</td>
<td>8</td>
<td>13.7</td>
</tr>
<tr>
<td>Schleswig-Holstein</td>
<td>5</td>
<td>8</td>
<td>13.7</td>
</tr>
<tr>
<td>Thüringen</td>
<td>3</td>
<td>8</td>
<td>13.7</td>
</tr>
</tbody>
</table>

### Table 7: Fixed-effects regressions at the county level. Dependent variables are the overall number of childcare facilities, the number of facilities providing care to school-aged children, the number of places provided by and the number of employees in facilities.

<table>
<thead>
<tr>
<th></th>
<th>Facilities overall</th>
<th>Facilities for school-aged children</th>
<th>Places</th>
<th>Persons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Share ADS</td>
<td>-1.357 (3.634)</td>
<td>-0.967 (0.724)</td>
<td>112.129 (248.386)</td>
<td>-40.217 (85.639)</td>
</tr>
<tr>
<td>GDP p. c.</td>
<td>0.000 (0.000)</td>
<td>-0.000 (0.000)</td>
<td>0.000 (0.005)</td>
<td>0.000 (0.002)</td>
</tr>
<tr>
<td>Unemployment rate</td>
<td>113.540** (52.709)</td>
<td>32.179** (14.884)</td>
<td>12555.605*** (3660.303)</td>
<td>2874.635*** (1067.536)</td>
</tr>
<tr>
<td>Population density</td>
<td>0.227** (0.106)</td>
<td>-0.007 (0.008)</td>
<td>14.021** (6.535)</td>
<td>4.945*** (1.727)</td>
</tr>
<tr>
<td>Population share children</td>
<td>1563.236*** (371.682)</td>
<td>5.645 (57.269)</td>
<td>1.14×10^7 *** (2.2×10^6)</td>
<td>29095.870*** (6869.301)</td>
</tr>
</tbody>
</table>

Year dummies yes
State × year dummies yes
R² 0.480
N 2547

Fixed-effects regressions at the county level for the time range 2007–2014. Heteroskedasticity-robust standard errors in parenthesis. ∗/∗∗/∗∗∗ indicate significance at the 10%/5%/1%-level. † provided by/working in all childcare facilities, no restriction by age of children. ‡ Here, children refers to those under the age of ten. Data sources are listed in section 11.

<table>
<thead>
<tr>
<th>Categories</th>
<th>Facilities overall</th>
<th>Facilities for school-aged children</th>
<th>Places</th>
<th>Persons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Share ADS</td>
<td>-1.357 (3.634)</td>
<td>-0.967 (0.724)</td>
<td>112.129 (248.386)</td>
<td>-40.217 (85.639)</td>
</tr>
<tr>
<td>GDP p. c.</td>
<td>0.000 (0.000)</td>
<td>-0.000 (0.000)</td>
<td>0.000 (0.005)</td>
<td>0.000 (0.002)</td>
</tr>
<tr>
<td>Unemployment rate</td>
<td>113.540** (52.709)</td>
<td>32.179** (14.884)</td>
<td>12555.605*** (3660.303)</td>
<td>2874.635*** (1067.536)</td>
</tr>
<tr>
<td>Population density</td>
<td>0.227** (0.106)</td>
<td>-0.007 (0.008)</td>
<td>14.021** (6.535)</td>
<td>4.945*** (1.727)</td>
</tr>
<tr>
<td>Population share children</td>
<td>1563.236*** (371.682)</td>
<td>5.645 (57.269)</td>
<td>1.14×10^7 *** (2.2×10^6)</td>
<td>29095.870*** (6869.301)</td>
</tr>
</tbody>
</table>

Year dummies yes
State × year dummies yes
R² 0.480
N 2547

Fixed-effects regressions at the county level for the time range 2007–2014. Heteroskedasticity-robust standard errors in parenthesis. ∗/∗∗/∗∗∗ indicate significance at the 10%/5%/1%-level. † provided by/working in all childcare facilities, no restriction by age of children. ‡ Here, children refers to those under the age of ten. Data sources are listed in section 11.

<table>
<thead>
<tr>
<th>Categories</th>
<th>Facilities overall</th>
<th>Facilities for school-aged children</th>
<th>Places</th>
<th>Persons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Share ADS</td>
<td>-1.357 (3.634)</td>
<td>-0.967 (0.724)</td>
<td>112.129 (248.386)</td>
<td>-40.217 (85.639)</td>
</tr>
<tr>
<td>GDP p. c.</td>
<td>0.000 (0.000)</td>
<td>-0.000 (0.000)</td>
<td>0.000 (0.005)</td>
<td>0.000 (0.002)</td>
</tr>
<tr>
<td>Unemployment rate</td>
<td>113.540** (52.709)</td>
<td>32.179** (14.884)</td>
<td>12555.605*** (3660.303)</td>
<td>2874.635*** (1067.536)</td>
</tr>
<tr>
<td>Population density</td>
<td>0.227** (0.106)</td>
<td>-0.007 (0.008)</td>
<td>14.021** (6.535)</td>
<td>4.945*** (1.727)</td>
</tr>
<tr>
<td>Population share children</td>
<td>1563.236*** (371.682)</td>
<td>5.645 (57.269)</td>
<td>1.14×10^7 *** (2.2×10^6)</td>
<td>29095.870*** (6869.301)</td>
</tr>
</tbody>
</table>

Year dummies yes
State × year dummies yes
R² 0.480
N 2547

Fixed-effects regressions at the county level for the time range 2007–2014. Heteroskedasticity-robust standard errors in parenthesis. ∗/∗∗/∗∗∗ indicate significance at the 10%/5%/1%-level. † provided by/working in all childcare facilities, no restriction by age of children. ‡ Here, children refers to those under the age of ten. Data sources are listed in section 11.

<table>
<thead>
<tr>
<th>Categories</th>
<th>Facilities overall</th>
<th>Facilities for school-aged children</th>
<th>Places</th>
<th>Persons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Share ADS</td>
<td>-1.357 (3.634)</td>
<td>-0.967 (0.724)</td>
<td>112.129 (248.386)</td>
<td>-40.217 (85.639)</td>
</tr>
<tr>
<td>GDP p. c.</td>
<td>0.000 (0.000)</td>
<td>-0.000 (0.000)</td>
<td>0.000 (0.005)</td>
<td>0.000 (0.002)</td>
</tr>
<tr>
<td>Unemployment rate</td>
<td>113.540** (52.709)</td>
<td>32.179** (14.884)</td>
<td>12555.605*** (3660.303)</td>
<td>2874.635*** (1067.536)</td>
</tr>
<tr>
<td>Population density</td>
<td>0.227** (0.106)</td>
<td>-0.007 (0.008)</td>
<td>14.021** (6.535)</td>
<td>4.945*** (1.727)</td>
</tr>
<tr>
<td>Population share children</td>
<td>1563.236*** (371.682)</td>
<td>5.645 (57.269)</td>
<td>1.14×10^7 *** (2.2×10^6)</td>
<td>29095.870*** (6869.301)</td>
</tr>
</tbody>
</table>

Year dummies yes
State × year dummies yes
R² 0.480
N 2547

Fixed-effects regressions at the county level for the time range 2007–2014. Heteroskedasticity-robust standard errors in parenthesis. ∗/∗∗/∗∗∗ indicate significance at the 10%/5%/1%-level. † provided by/working in all childcare facilities, no restriction by age of children. ‡ Here, children refers to those under the age of ten. Data sources are listed in section 11.
<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Minimum</th>
<th>95th Percentile</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mother-/Household-level, N = 17696</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hourly Wage (incl. imputations)</td>
<td>13</td>
<td>7.2</td>
<td>0.02</td>
<td>25</td>
<td>186</td>
</tr>
<tr>
<td>Monthly Nonlabor Income</td>
<td>4056</td>
<td>19039</td>
<td>0</td>
<td>11681</td>
<td>2110083</td>
</tr>
<tr>
<td>Migration Background†</td>
<td>0.31</td>
<td>0.46</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>No. Children Aged &lt; 1</td>
<td>0.02</td>
<td>0.13</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>No. Children Aged 1–2</td>
<td>0.13</td>
<td>0.35</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>No. Children Aged 3–5</td>
<td>0.21</td>
<td>0.43</td>
<td>0</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>No. Children Aged 12–15</td>
<td>0.35</td>
<td>0.57</td>
<td>0</td>
<td>1</td>
<td>3</td>
</tr>
</tbody>
</table>

† Binary variable, therefore only mean and standard deviation given

**Table 8: Additional Summary Statistics of the SOEP data on mothers/households**

<table>
<thead>
<tr>
<th>Federal State</th>
<th>N</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baden-Württemberg</td>
<td>3648</td>
<td>14.79</td>
</tr>
<tr>
<td>Bavaria</td>
<td>5171</td>
<td>20.97</td>
</tr>
<tr>
<td>Berlin</td>
<td>910</td>
<td>3.69</td>
</tr>
<tr>
<td>Brandenburg</td>
<td>1047</td>
<td>4.25</td>
</tr>
<tr>
<td>Hamburg</td>
<td>450</td>
<td>1.82</td>
</tr>
<tr>
<td>Hesse</td>
<td>2342</td>
<td>9.50</td>
</tr>
<tr>
<td>North Rhine-Westphalia</td>
<td>6934</td>
<td>28.11</td>
</tr>
<tr>
<td>Rhineland-Palatinate</td>
<td>1536</td>
<td>6.23</td>
</tr>
<tr>
<td>Saarland</td>
<td>231</td>
<td>0.94</td>
</tr>
<tr>
<td>Schleswig-Holstein</td>
<td>1240</td>
<td>5.03</td>
</tr>
<tr>
<td>Thuringia</td>
<td>1154</td>
<td>4.68</td>
</tr>
<tr>
<td><strong>Σ</strong></td>
<td>24663</td>
<td>100</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year</th>
<th>N</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002</td>
<td>207</td>
<td>0.84</td>
</tr>
<tr>
<td>2003</td>
<td>221</td>
<td>0.90</td>
</tr>
<tr>
<td>2004</td>
<td>1088</td>
<td>4.41</td>
</tr>
<tr>
<td>2005</td>
<td>1035</td>
<td>4.20</td>
</tr>
<tr>
<td>2006</td>
<td>1211</td>
<td>4.91</td>
</tr>
<tr>
<td>2007</td>
<td>1346</td>
<td>5.46</td>
</tr>
<tr>
<td>2008</td>
<td>1180</td>
<td>4.78</td>
</tr>
<tr>
<td>2009</td>
<td>1153</td>
<td>4.68</td>
</tr>
<tr>
<td>2010</td>
<td>2900</td>
<td>11.76</td>
</tr>
<tr>
<td>2011</td>
<td>3466</td>
<td>14.05</td>
</tr>
<tr>
<td>2012</td>
<td>3475</td>
<td>14.09</td>
</tr>
<tr>
<td>2013</td>
<td>4009</td>
<td>16.26</td>
</tr>
<tr>
<td>2014</td>
<td>3372</td>
<td>13.67</td>
</tr>
<tr>
<td><strong>Σ</strong></td>
<td>24663</td>
<td>100</td>
</tr>
</tbody>
</table>

**Table 9: Observations (at child-level) per federal state and year**
11.3 County-Level Data

To control for county characteristics data from the databases of the federal and state statistical offices is used.

Population density is computed using counties’ land area (Statistische Ämter des Bundes und der Länder 2016e) and population sizes (Statistische Ämter des Bundes und der Länder 2016c).

Unemployment rates are supplied by Statistische Ämter des Bundes und der Länder (2016b).

Data on the per capita GDP in the counties is mostly drawn directly from Statistische Ämter des Bundes und der Länder (2016d) which, however, contains a number of missing values for the years 2013 and 2014. These missings were imputed using the last available county-level value of GPD p.c. and the state-level growth rate from Arbeitskreis "Volkswirtschaftliche Gesamtrechnungen der Länder" (2016). Data on childcare facilities other than all-day schools are drawn from Statistische Ämter des Bundes und der Länder (2016f).

Table 10 lists summary statistics of the county-level variables.

12 Estimation Results Including Covariates

See table 11.

13 Calculation of Fiscal Effects

This calculation assumes that the findings of section 5.3 generalize to Germany as a whole, i.e. that the share of students visiting all-day schools increases by 17.6 pp and that their mothers work 13.72 weekly hours more (see table 4).

In 2014 2,708,752 children were in primary school, the according increase is 0.176 × 2,708,752 = 476,740. In this paper’s estimation sample the average mother has 1.39 children, thus 342,978 mothers are affected. As all-day schools are assumed to only affect mothers whose children were so far not in full-day care (thus, those who might switch from other means of childcare are ignored), their sample means are used. The average mother works 22.9 hours per week earning 13.63 €. With four weeks per month and assuming a constant hourly wage\(^{21}\) their yearly earnings change from

\[
22.9 \times 4 \times 12 \times 13.63 = 14,982 \text{€}
\]

to

\[
(22.9 + 13.72) \times 4 \times 12 \times 13.63 = 23,958 \text{€}.
\]

This yearly earning are used as total yearly income of single mothers. 84.6% of mothers are married, their monthly nonlabor-income amounts to on average 4,425 €. This is used to approximate their husbands’ income to account for the effects of income tax splitting. Thus, yearly household incomes of married mothers change from

\[
14,982 \times 12 \times 4425 = 680,825 \text{€}
\]

to

\[
23,958 \times 12 \times 4425 = 770,585 \text{€}.
\]

Tax payments are calculated\(^{22}\) according to §32a Einkommensteuergesetz (ESTG) (2014), with additional 5.5% Solidaritätszuschlag according to Solidaritätszuschlaggesetz 1995 (SolzG 1995) (2016) (reference year is 2014). The yearly change in tax payments therefore amounts to

\[
3941.48 - 1411.2 = 2530.28 \text{€}
\]

for singles and

\[
17753.54 - 14495.7 = 3257.84 \text{€}
\]

for married couples.

The expected additional tax revenue per mother is

\[
0.846 \times 3257.84 + 0.154 \times 2530.28 = 3145.79 \text{€}.
\]

Summed up over all mother this amounts to

\[
3145.79 \times 342978 = 1,078,936,762 \text{€}.
\]

\(^{21}\) As with an increase in hours the wage likely increases this approach computes a lower boundary.

\(^{22}\) and confirmed using an online service of the German Ministry of Finance: https://www.bmf-steuerrechner.de/ekst/
<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Minimum</th>
<th>95th Percentile</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>GDP p. c.</strong></td>
<td>30130</td>
<td>12926</td>
<td>12318</td>
<td>56831</td>
<td>126149</td>
</tr>
<tr>
<td><strong>Unemployment Rate</strong></td>
<td>0.074</td>
<td>0.039</td>
<td>0.012</td>
<td>0.15</td>
<td>0.25</td>
</tr>
<tr>
<td><strong>Population Density</strong></td>
<td>576</td>
<td>746</td>
<td>36</td>
<td>2296</td>
<td>4601</td>
</tr>
</tbody>
</table>

No. of Observations: 2986

**Table 10:** Summary statistics of county-level data

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Minimum</th>
<th>95th Percentile</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>GDP p. c.</strong></td>
<td>30130</td>
<td>12926</td>
<td>12318</td>
<td>56831</td>
<td>126149</td>
</tr>
<tr>
<td><strong>Unemployment Rate</strong></td>
<td>0.074</td>
<td>0.039</td>
<td>0.012</td>
<td>0.15</td>
<td>0.25</td>
</tr>
<tr>
<td><strong>Population Density</strong></td>
<td>576</td>
<td>746</td>
<td>36</td>
<td>2296</td>
<td>4601</td>
</tr>
</tbody>
</table>

**Table 11:** Instrumental variable estimations of the effect of full-day childcare (instrumented by the share of all-day schools in a county, Share ADS) on maternal labor supply including covariates. Labor supply is measure at the intensive (weekly working hours) and extensive margin (indicator for full- or parttime work). Results are given for the full, unrestricted sample. Estimation equations are set up corresponding to equation [10]. Columns I and E correspond to columns I.3 and E.3 in table 3.