

Lauber, Verena; Thomas, Lampert

**Conference Paper**

## The Effect of Early Universal Daycare on Child Weight Problems

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2014: Evidenzbasierte Wirtschaftspolitik  
- Session: Child Care, No. C02-V3

**Provided in Cooperation with:**

Verein für Socialpolitik / German Economic Association

*Suggested Citation:* Lauber, Verena; Thomas, Lampert (2014) : The Effect of Early Universal Daycare on Child Weight Problems, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2014: Evidenzbasierte Wirtschaftspolitik - Session: Child Care, No. C02-V3, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft, Kiel und Hamburg

This Version is available at:

<https://hdl.handle.net/10419/100399>

**Standard-Nutzungsbedingungen:**

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

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

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

**Terms of use:**

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

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

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

# The Effect of Early Universal Daycare on Child Weight Problems

**Verena Lauber\***  
University of Konstanz

**Thomas Lampert**  
Robert Koch-Institute

5th September 2014

*Preliminary version - Please do not cite without author's permission- Comments welcome*

## Abstract

As the number of young children in daycare increases, people start to worry about the effect of early non-parental care. This is of special relevance as investments in the early periods of life are shown to be most important for a child's long term development. Based on the German national health survey for children, we study the impact of daycare in pre-*kindergarten*-age on weight problems and gross motor skills of children aged five to nine. This dataset has the advantage to provide objective child development measures. Our results are thus not prone to reporting bias. We estimate the effect of early daycare based on a non-linear instrumental variable strategy, by exploiting regional differences in subsidized center-based care for zero to three year old children as the source of exogenous variation. Our OLS estimates reveal only very weak differences and do not hint at an increase in weight problems due to early use of daycare. The estimated local average treatment effects even indicate that early daycare leads to more desirable physical development for children 'at the margin'. Further analyses suggest that we do not estimate an effect for the most advantaged children, as the caring decision in families with low and medium income and with an overweight father seem to react most strongly to differences in daycare supply.

JEL Classification: I12, J13, I20, J18

Keywords: Childcare, obesity, fitness, non-linear instrumental variables

**Acknowledgments:** We thank Friedrich Breyer, C. Katharina Spieß, Rafael Lalive and the researchers in the department "Education and Family" at DIW Berlin for valuable comments. Comments of the participants of the Doctoral Seminar in Public Economics at the University of Konstanz, the Brown Bag Seminar at the University of Konstanz, the 10th TEPP conference in Le Mains, the 6th DGGÖ conference in Munich, the Workshop on Child Health and Education at the University of Surrey, the RWI Research Network Conference on the Economics of Education, the 5th International Workshop on Applied Economics of Education in Catanzaro and the 10th World Congress in Health Economics (iHEA & ECHE) in Dublin are also gratefully acknowledged. In particular we thank Ronny Kuhnert and Rüdiger Dölle of the Robert Koch-Institute for their help.

---

\*Corresponding Author. E-Mail: verena.lauber@uni-konstanz.de. Department of Economics, University of Konstanz.

# 1 Motivation

The increasing use of non-parental care for young children is not viewed uncritically. Non-parental care alternatives are crucial for enabling both parents to participate at the labor market, which seems to be the wish of the majority of parents<sup>1</sup>, but childcare arrangements have also the potential to influence child development, because they alter the composition of essential inputs a child receives early in life.

There exist both, voices claiming that non-parental care is harmful for child development and voices emphasizing it's benefits. Existing research shows that daycare, under certain circumstances, can have a positive impact on the child development in particular for disadvantaged children. High quality interventions are shown to be beneficial, even in the long-run.<sup>2</sup> However, this research does not by far cover all relevant aspects of child development. Physical and health development is greatly disregarded even though in particular weight problems have become a global epidemic, which also affects children at an increasing rate.<sup>3</sup>

Daycare, referring to non-parental care provided by care centers or trained private persons as childminders, may actually have a good chance to improve the child's physical development. Children may for example be more active and their eating behavior well-regulated. There is evidence that dietary intake and physical activity account for more in the variance in changes in a child's Body-Mass-Index than parental obesity (Klesges et al. 1995). Eating habits and food preferences also seem to develop very early in life (Birch 1999), thus having a life-long impact.

Child weight problems and reduced fitness are related to a myriad of health consequences, such as hypertension and other cardiovascular disease risk factors, type 2 diabetes and obstructive sleep apnea (Ebbeling et al. 2002). Private and social cost are considered to be even higher as childhood obesity is strongly associated with adult weight problems (Serdula et al. 1993; Freedman et al. 2005; De Onis et al. 2010). The health-related cost of overweight and obesity, by increasing the need for medical treatment create a non-negligible burden to the health care systems in the US as well as in Europe (see e.g. Fry and Finley 2005; Konnopka et al. 2011; Tsai et al. 2011).<sup>4</sup> Research, however, reveals additional costs of child weight problems. Childhood obesity not just hinders a child's skill attainment (Cawley and Spiess 2008), but bad childhood health also seems to affect future education and income (e.g. Currie 2009; Currie et al. 2010; Conti et al. 2010).<sup>5</sup> Preventing weight and fitness problems has consequently high social benefits.

Investments early in life are in general shown to be more productive than later investments. Preventing bad

---

<sup>1</sup>Even in the 1990s the male-breadwinner model has been shown to be the least preferred employment pattern in most OECD countries (OECD 2001: chapter 4). The increasing level of education among women, the awareness of the role of disruption and reduced working hours for wages, the challenges of demographic change and the shortages of skilled labor even made female labor market participation an explicit political goal in countries such as Germany (see e.g. OECD 2014).

<sup>2</sup>(See e.g. Heckman et al. 2010; Chetty et al. 2011; Havnes and Mogstad 2011).

<sup>3</sup>(See e.g. Finucane et al. 2011; Ng et al. 2014).

<sup>4</sup>Konnopka et al. (2011) estimate obesity to causes €4,854 million in direct costs, corresponding to 2.1 percent of the overall German health expenditures in 2002 and €5,019 million EUR in indirect costs.

<sup>5</sup>There are also direct negative effect on wages (Cawley 2004). Following psychological diseases, muscle and heart problems are also major causes of occupational disability (Deutsche Rentenversicherung 2012).

development early seems to be more efficient than paying for the treatment of the consequences later on (Heckman and Masterov 2007). It is argued these early investments are in particular important, because of self-productivity and dynamic complementary to investments later in life (Cunha and Heckman 2007).<sup>6</sup> There is also evidence directly related to obesity confirming the strong effect of early interventions (Davis and Christoffel 1994). Different childcare settings can be seen as diverging investments early in life and as such be expected to make a difference for child development.

It is theoretically undetermined if substituting parental (caring) time for daycare has a positive or negative impact on child development. If daycare is a good substitute for parental care or even provides an input parents cannot to provide (e.g. playing/exercising with other children) one would expect daycare to have a positive impact. This effect could be reinforced if families change to a dual earner household, with all the attendant benefits of increased household income.<sup>7</sup> If, however, parental time is an essential factor in the child quality production function or if the marginal productivity of daycare is low, parents may harm the child by relying on daycare.<sup>8</sup> One should keep in mind that daycare does not fully substitute for parental care, it may serve as a complement and as such be valuable at least for some children.

Evidence on the effect of daycare on physical development is still very limited. A larger part of the literature addresses cognitive and behavioral development. Existing evidences on the effect of daycare on child physical health is pre-dominantly derived from targeted programs in the US.<sup>9</sup> The German daycare system can, however, be described as universal, as it is not targeted at any specific group. It is not the primary business of these institutions to support disadvantaged children. They rather have an educational mission approaching children independent of family background.

In this paper we exploit a unique dataset on child health, the national health interview and examination for children and adolescents ‘*Studie zur Gesundheit von Kindern und Jugendlichen in Deutschland*’ (KiGGS) conducted in Germany in the years 2003 to 2006 to estimate the effect of early daycare usage on child weight and fitness development. This dataset is representative for Germany and has the advantage that objective health measures and tests are conducted. This is important, because there exists evidence that health measures, especially related to weight, reported by parents or children themselves, can be relatively poor measures of the actual status (Goodman et al. 2000). Parents who are aware of the optimal child develop-

---

<sup>6</sup>The idea that investments early in life when the brain is more plastic are most efficient is also supported by neurologists and psychiatrists (see e.g. Knudsen et al. 2006).

<sup>7</sup>Blau (1999) and Currie and Lin (2007) give extensive overviews over the existing literature on the child health income gradient. In addition maternal employment could increase maternal life satisfaction, which in turn could have a positive impact on child development (Berger 2013).

<sup>8</sup>Based on the ideas of Becker (1965), that each commodity that spends utility is produced using market and time inputs, and Grossman (1972), that current health depends on past health status, as well as medical inputs, the reduced form production function of interest can be illustrated as:  $H = H(T, M, X)$ , where  $T$  refers to parental time investments and  $M$  to market goods, which in this case includes daycare. An increase in working hours potentially increases daycare usage, if it is not fully compensated by a reduction of other time use.  $X$  may include parental and child characteristics, that are often unobservable, but relevant for child development.

<sup>9</sup>Child development is multi-dimensional. A large part of the literature on the effects of early education deals with cognitive and non-cognitive development. A small part focuses on health in terms of infectious illnesses. In a survey D’Onise et al. (2010) summarize that early educational institutions have relatively good chances to reduce weight problems.

ment may overstate their child's health. Related to childcare it seems plausible that parental reporting is correlated with the choice of the childcare mode.<sup>10</sup>

Burkhauser and Cawley (2008) recommend that one indicator is not enough to determine weight problems, therefore in the present study the traditional measurement of weight problems based the Body-Mass-Index is supplemented by a measure of the percentage of body fat. In addition the performance in a test for gross motor skills is analyzed. Gross motor skills can be viewed as a complementing measure of future physical activity and as such health risk (Stodden et al. 2009; D'Hondt et al. 2009; Barnett et al. 2008).

We focus on the estimation of the effect of early daycare usage, starting in pre-*kindergarten*-age. We assume that we estimate an effect of relatively high quality care as the government in Germany regulates and monitors the quality of public daycare (Spieß and Tietze 2002). The quality of German daycare was not criticized very heavily compared to other countries as for example the US.<sup>11</sup> We also do not analyze the effect of a strong expansion in daycare supply, which has in some cases shown to be associated with a worsening of child well-being, potentially because of a decrease in quality (Baker et al. 2008).<sup>12</sup>

The identification of a causal effect of early daycare is not a simple task. Parents try to contrast cost and benefits. In this calculation not only does the child's development play a role, but also the parents' personal objectives and in particular the mother's desire to participate in the labor market. Their attitudes towards caring and their desire for spending time with children are also crucial. The precision of the parental assessment of costs and benefits may vary with parental characteristics and it is based on information that is often not observable to the researcher. A simple comparison of the development of children with and without early daycare may be biased, because of sorting and selection problems.

We address the identification problem by exploiting regional differences in the supply with slots in subsidized center-based care across German counties. For children younger than three years of age the supply of subsidized daycare is rather limited, with less than three slots per 100 children available in some areas over the observation period.<sup>13</sup> There are a few private providers of regular care, mainly childminders, but the market is very underdeveloped (Kreyenfeld and Hank 2000). There is evidence that the demand strongly exceeds the supply. This is the case both in West and East Germany, with East Germany having a higher supply for historical reasons (Wrohlich 2008). For children, aged three to six, a daycare slot in *kindergarten* is ensured by law since 1996 and the provision is at nearly 90 percent.<sup>14</sup> A higher supply of subsidized daycare for children younger than three years sets incentives for parents to start using daycare before the child turns three years old, and as such would receive a slot in *kindergarten*, anyway.

---

<sup>10</sup>Even though if it is not clear whether the relationship would be positive or negative.

<sup>11</sup>It is acknowledged that the quality of the substitute for family care matters for the demand (Blau 2001) and the outcome (Datta Gupta and Simonsen 2010; Blau and Currie 2006).

<sup>12</sup>Our estimates may also provide a measure of part-time daycare (Spieß et al. 2002). If mothers work in Germany they most often work part-time (Renz and Eggen 2004).

<sup>13</sup>For these children families are the main care takers. The role of center-based care increases with age (Spieß et al. 2002).

<sup>14</sup>Only in the last years the government exhibited strong efforts to increase public daycare provision, because from August 2013 on a daycare slot for children that are younger than three years old is also guaranteed by law.

We cope with the binary nature of the treatment by implementing a non-linear instrumental variable strategy. We perform a three-step procedure and as such exploit the predicted probability to use daycare early as the actual instrument. This instrument has been shown to be more efficient and to lead to more precisely estimated effects (e.g. Mogstad and Wiswall 2012). The strategy is also robust to an incorrect specification of the non-linear model for the participation decision and does not require the adjustment of the usual two-stage-least-squares standard errors for statistical inference (Wooldridge 2002).

This paper contributes to the literature by focusing on the estimation of a causal effect in a setting with two factors that are rarely analyzed together: Early universal daycare and child weight and fitness development. The analysis takes part in Germany, a country with limited evidence on the effects of its daycare system. Based on the linear OLS estimates we can state, that daycare usage with 30 months or earlier is not associated with higher risk of overweight and obesity on average. For lower and middle income groups we even find estimates, that hint to fewer weight problems in the treatment group. At the same time physical fitness, measured by the gross motor skills test seems to be better. The non-linear instrumental variable estimation leads to the conclusion that children ‘at the margin’ benefit from early daycare usage in terms of significantly fewer weight problems and a better performance in the gross motor skills test. The instrument creates the strongest variation in the group of children in non-high income households and families with an overweight father (Body-Mass-Index > 25). As such we estimate effects for the less advantaged children.

The remainder of this paper is divided in five sections. Section 2 gives an overview over the related literature. Section 3 describes the dataset and the estimation procedure. Section 4 introduces the instrument. Section 5 provides the results on the effect of early daycare usage on child weight and fitness development and discusses the sensitivity of the results to the chosen specification. Section 6 concludes.

## 2 Related Literature

There exists a literature on the effect of care arrangements on child development, but there is very little on its effect on health outcomes, in particular related to weight problems and physical fitness. If there are studies related to health, these can mostly be found in the US. Childcare institutions, however, are country-specific and vary in terms of quality and form. There is a larger body of research on the effects of non-parental care in *kindergarten*-age of three to six. In contrast to this, the the research on the effects of daycare during earlier ages is fairly small.<sup>15</sup>

One part of this literature defines the ‘care treatment’ simply as differences in maternal time available for children and as such focuses on maternal employment. The evidence on child physical and weight development based on this definition is by far the largest. However, it has to be kept in mind, that maternal

---

<sup>15</sup>Parents who decide to use daycare early may also differ in terms of their attitudes and norms. The evidence on the effect of early daycare is difficult to compare. The group of children under consideration also varies essentially. Few studies focus on the very young children, aged three or younger, as it is the case in the present study. For many countries and outcomes there is a lack of evidence.

employment is not the direct counterpart of daycare usage. Often we do not know, who is responsible for the child while the parents are working. It is, consequently, not surprising that the results depend on the working intensity, the quality of the substitute for parental care and the age of the children. For Germany Dustmann et al. (2012) find for example that there was no improvement in child development after an increase of the maternal time by a lengthening of the maternal leave period<sup>16</sup>, while Baum II (2003) for the US shows that maternal employment in the first year of a child's life has a negative impact on white children's cognitive and behavioral development. Mahler (2008) analyzes the effect of maternal labor supply on weight problems in Germany. He finds that a high working intensity, in terms of the aggregate number of years the mother worked full-time until the child was 15 years old, increases the risk to become obese for young adults, but there is no effect for part-time work. Part-time work and as such part-time care is, however, much more common for mothers in Germany.<sup>17</sup> For the US Lumeng et al. (2004) find that some center-based daycare attendance, in terms of hours per week, compared to no center-based care attendance is associated to a decreased risk of being overweight.<sup>18</sup> Even if we do not directly focus on the caring intensity, we take away from these findings that it is plausible to find positive effects in Germany, where hours in daycare are not overly high on average (see e.g. Spieß et al. 2002).

The majority of studies on maternal employment and child physical health development reveals a negative relationship.<sup>19</sup> Mindlin et al. (2009) for example compares several OECD countries and finds a positive correlation of maternal employment and vaccination status, but also with the risk of obesity.<sup>20</sup> However, some studies also show that the effects depend on the child's family background. The increase in weight problems is in particular driven by children with high socio-economic background (Anderson et al. 2003; Ruhm 2008).<sup>21</sup>

---

<sup>16</sup>Even though it had an impact on the return to work for mothers.

<sup>17</sup>Germany is a country with a tradition of (long) paid leave periods and a high fraction of non-working or part-time working mothers. Maternal full-time work is a rather uncommon phenomenon at least in households with young children in Germany (Renz and Eggen 2004). In the observation period 60 percent of parents report to use daycare only part-time (Spieß et al. 2002). It is even likely that the author does not capture all unobservable factors driving mothers into full-time work, even if he has a large set of controls variables.

<sup>18</sup>Even extensive usage is not associated with an increase in future overweight. Directly focusing on daycare usage, Datta Gupta and Simonsen (2010) show for 11 years old children in Denmark that daycare by more than 30 hours per week could be harmful for socio-emotional development, but fewer hours do not have a negative effect. The starting age also seems to matter. For the US, Loeb et al. (2007) find based on an OLS estimation that a starting age between two and three years is associated with the highest academic benefits, compared to earlier and later starting ages. However, behavioral problems are stronger if daycare is used very early.

<sup>19</sup>Most US studies find negative effects of early maternal employment on all kind of child development indicators, including cognitive and behavioral development, immunization, medical check-ups and infant mortality (Ruhm 2000; 2004; Berger et al. 2005; Tanaka 2005; Rossin 2011).

<sup>20</sup>Some more studies in economic and medical sciences that deal with the relationship between maternal employment and obesity confirm negative or mixed effects. von Hinke Kessler Scholder (2008) reveals a significant positive correlation between full-time maternal employment during mid-childhood, but not during early or later childhood, and the probability of being overweight of children at the age of 16 in the United Kingdom. Classen and Hokayem (2005) focus on two- to 18-year-old American children, Phipps et al. (2006) analyzes six to 11 year old children in Canada, Cawley and Liu (2012) focuses on Spanish children aged two to 15 and Takahashi et al. (1999) on three-year-old Japanese children. They all show a positive tendency on the risk to become obese. Benson and Mokhtari (2011) reveal that in the US also the father's employment has a negative but increasing effect on child weight problems.

<sup>21</sup>Similarly, but not with the focus on health outcomes, Waldfogel et al. (2002) and del Carmen Huerta et al. (2011) find for the US and several other OECD countries find that for disadvantaged groups the income effect may offset potential negative effects.

A second crucial determinant seems to be the quality of the substitute for parental care, which is relatively highly regulated on Germany (see Spieß and Tietze 2002). Herbst and Tekin (2010; 2011; 2012) for example, analyzing children of single disadvantaged mothers in the US, reveal an increase in weight problems and negative effects on reading and math skills if the mother is a childcare subsidy recipient.<sup>22</sup> Herbst and Tekin (2011) argue that low quality center-based care is the a key mechanism.<sup>23</sup> Concurrently, Votruba-Drzal et al. (2004) find in a small US sample that high quality daycare is associated to fewer behavioral problems for two to four year olds children in particular with low-income background. Similarly, Datta Gupta and Simonsen (2010) show for Denmark that sons of mothers with vocational training gain if they were in pre-school instead of ‘family daycare’, which is described to be of relatively low quality.<sup>24</sup> Chetty et al. (2011) analyzing the project Star in the US highlight the importance of class size and teacher quality.

Hence it may not be surprising that the US-evidence, drawn from high quality interventions that target disadvantaged children, reveals mainly positive short and long-run effects (see e.g. Currie and Almond 2011; Heckman et al. 2010). It is obvious that in this case we cannot speak of universal childcare. This strand of research represents the majority of studies that directly deal with care institutions and not maternal employment. It is questionable how much we can learn from these findings about the effect of universal care, because these interventions are often relatively intensive and the targeted group is not representative. These programs also seem to be beneficial for child health development.<sup>25</sup> Frisvold (2006) and Carneiro and Ginja (2008) show that participation in *Head Start* reduces the obesity risk for children. Frisvold (2006) estimates mid- and long-run effects for five- to 19-year-old children based on a selection on observables and an exclusion strategy. Allowing for unobserved selection he finds strong a reduction in obesity prevalence of 28 percentage points among black children. Frisvold and Lumeng (2011) even find that full-day participation outperforms part-time care in *Head Start* in terms of obesity prevention. With their value-added approach they estimate a reduction of obesity prevalence by about 18 percentage points for attending full-day *Head Start*. In their reduced form as well as in the structural modeling approach Carneiro and Ginja (2008) find that *Head Start* reduces of the risk to be overweight for 12- to 13-year-old children by 20 percentage points

---

Lamontagne et al. (1998) also reveal a positive effect of maternal employment on child development in Nicaragua. This is mainly due to an increase in income.

<sup>22</sup>There are heterogeneous effects. Based on an IV strategy Herbst and Tekin (2012) show that children of low-skilled mother suffer, while those of high-skilled mother may even benefit from the childcare subsidy related to weight problems. Their hypothesis is that high skilled mothers get better paid and more flexible jobs, potentially with health insurance.

<sup>23</sup>In their subsample of single mothers the intensity to which maternal working hours increase may also play a crucial role. Maher et al. (2008) using the same US dataset confirm the positive association between daycare and obesity. Bernal and Keane (2011) on the other hand, estimating the effect of an increase in daycare usage with a simultaneous increase in working hours for single mothers via an instrumental variables strategy in the US, find no adverse effect of center-based care on cognitive achievement. They did not analyze health problems.

<sup>24</sup>There is, however, no large difference between pre-school and home care. For Germany Peter (2013) reveals that a low child-staff ratio seems underlie a positive effect of early daycare usage.

<sup>25</sup>The effects on child cognitive and school outcomes are mostly found to be positive for children of primary school age (Fitzpatrick 2008; Berlinski et al. 2009). Very targeted pre-school programs, as the *Perry Pre-School* program or *pre-kindergarten* programs in Oklahoma or Texas, have been shown to yield high positive returns related to various outcomes (Currie and Thomas 1995; Heckman et al. 2010; Gormley Jr et al. 2005; Andrews et al. 2012). Also slightly more universal programs, as *Head Start* have positive effects, in the short run (Heckman 2000) and the long run (Garces et al. 2002).

on average. In a recent study Campbell et al. (2014) are able to show that the *Carolina Abecedarian Project* has significantly reduced child and adult weight problems. These effects, however, only exist for males. They find a more than 40 percentage points reduction of the risk to be obese in the group of 9- to 12-months-old boys and a by 29 percentage points reduced overweight prevalence among eight-year old boys. The risk to suffer from obesity and hypertension was estimated to be half as high for men who if they were enrolled in the *Carolina Abecedarian Project*.

The evidence on the effects of universal care, as it is common in most European countries, is rather limited and to our knowledge there is no study focusing on the effect of this universal daycare for under three-year-old children on child weight and fitness problems.<sup>26</sup> Findings for non-health outcomes are mainly positive, apart from evidence in Canada. There, Baker et al. (2008) exploit the introduction of universal subsidized daycare in Quebec. They find that the increase in daycare usage had detrimental effect on parent-reported child outcomes, as motor skills, social skills and illnesses. At the same time it also decreases parental well-being. One explanation for this finding is again related to quality, as the strong and fast increase in supply may have reduce available resources per child (for that argument see Datta Gupta and Simonsen 2010).<sup>27</sup> Other studies reveal positive effects of universal care. Cascio (2009) provide evidence for a positive effect of an increase in state funded *kindergarten* programs for five-year-old children on school performance at least for white children in the US. Evidence also suggests an improvement of child development due to even earlier daycare usage. Havnes and Mogstad (2011) analyze a large scale expansion of childcare coverage for three- to six-year-old children in 1975 in Norway. Based on a difference-in-difference approach, they find positive long-term effects on education and labor market outcomes. For France, Dumas and Lefranc (2012) evaluate the effect of an increase in daycare provision, that strongly raised enrollment of three- and four-year old children. They find positive effects of the length of daycare (pre-school) experience on primary school performance. The lower and intermediate social groups seem to benefit the most. Exploring the effect of an increase in public daycare supply for three-year-old children in Spain, Felfe et al. (2013) find positive effects on math and reading skills of 15-year old adolescents. They find a particular positive effect for children from a disadvantaged background. Spain is an interesting case, because as in Germany subsidized daycare is expected to crowd out maternal care, rather than private care arrangements, as it could be the case for example in the US. Noboa-Hidalgo and Urzua (2012) analyze very early daycare attendance in the age of five to 14 months in Chile. They find positive effects on cognitive and motor skills development, that are most pronounced among children aged seven to twelve months. For the present study it is in particular interesting that gross motor skills and adaptive eating behavior were included in the study. Children gain

---

<sup>26</sup>In non-economic research as pediatrics the focus is more often on communicable illnesses (e.g. Bradley et al. 2003). Psychological research focuses more strongly on behavior and psychic outcomes, but also cognitive skills and language (see B auerlein et al. 2013: fro Germany).

<sup>27</sup>DeCicca and Smith (2013) even show that longer *kindergarten* duration in Canada leads to higher probabilities to repeat the third grade and decreases math and reading scores. However, in their case school entry age is affected and *kindergarten* seems to 'crowd out' school and other daycare.

related to both outcomes.<sup>28</sup> Even if gross motor skills are defined differently and eating here refers to caring out food-related tasks, these findings indicate that such factors can be influenced by daycare.

One of the few studies estimating the relationship between daycare and weight problems in the German daycare system, focuses on the hours in daycare for children aged three to five years (in *kindergarten*). Based on their regression adjustment analysis, Rapp et al. (2005) conclude that the obesity risk is not associated with the either being in part-time or full-time care.<sup>29</sup> The study by Felfe and Lalive (2013) is presumably closest related to the present one. They use a sample on the German school readiness test conducted in Schleswig-Holstein from 2009 to 2012 to estimate the effect of pre-*kindergarten* daycare on the evaluation of conspicuous development in several areas. All indicators suggest in agreement, that there is a reduction of problematic development if the child was in daycare earlier. The analysis also includes the assessment of motor skill development. However, this evaluation does not allow conclusions about physical development in the spirit of the present paper, as it was heavily based on fine motor skills.<sup>30</sup> We, consequently, focus on a relationship that has been under-researched so far.

### 3 Data and Estimation Strategy

The KiGGS data is a survey conducted by the Robert Koch Institute (RKI) on the health of children and adolescents in Germany. The study supplies data, representative of the country as a whole. The analysis in the present paper is based on version 4 of the baseline study of the years 2003 to 2006 (Kurth et al. 2013). 17641 children and adolescents aged between zero and 17 years participated in 167 sample points across all states of the German Federation. They filled in a questionnaire themselves if they were at least 11 years old. Otherwise the parents resumed this task. In addition parents provided information on their own background and the child's development since birth. This information was completed by an interview, an examinations and tests conducted by a trained person.

In the KiGGS, parents were asked if their child is or was cared only by the family before school enrollment. If the answer was "no", they were asked how old the child was when this non-family care was introduced. In addition parents report the type of non-family care: Either center-based care or care by a childminder<sup>31</sup>. In line with the official definition in Germany both types of care belong to formal care. We will stick to the term daycare, when we refer to these two types of care in the following. Unfortunately, we only have information about the starting age, but not about the hours in daycare.

---

<sup>28</sup>Even is significance is weak.

<sup>29</sup>This is in line with other research that revealed no strong association between *kindergarten* attendance and cognitive development in terms of later school placement in Germany (Spiess et al. 2003). Interventions that happen relatively close to school enrollment have also been shown to have very small effects in other countries as for example Norway (Drange et al. 2012).

<sup>30</sup>In a different version of their paper based on the German Socio-Economic Panel they do only find significant positive effects on social skills, but not on all motor, language and daily skills indicators (Felfe and Lalive 2012). However, there were no negative effects.

<sup>31</sup>Childminders are trained persons who most often care for a small group of children at home

We include all children for whom the parents report that they were only in center-based care or in center-based care and in care of a childminder. In the main specification we do not include children, for whom parents report that a childminder was the only type of care. In Germany childminders are mainly employed for relatively young children. Normally, children change to center-based care at latest when they turn three, because a slot is ensured by law from thereon. If a child was only in care of a childminder, we have to assume that non-family care was not permanent, such that it is less appropriate to interpret our threshold as a starting age. However, of all children surveyed in the relevant age-group, less than two percent were only in care of a childminder (see table A12).

We estimate the effect of pre-*kindergarten* daycare. In Germany children traditionally do not start to be in daycare much before the age of three, which we refer to as *kindergarten*-age, because institutions for these children are called *kindergartens*. For children aged three to six years a slot in such an institution is guaranteed by law since 1996, such on average nearly 90 percent of children are in daycare (DJI 2008).<sup>32</sup> The supply and usage of subsidized center-based care for up to three year-old children is much lower than for older children (see section 4). Figure 1 pictures the distribution of the starting age for daycare in our sample. It is obvious that there is a strong peak at the age of three, with a few individuals already starting one or two months earlier.

Unfortunately, only around two thirds of the parents report the starting age in month, all others just report years. We assume that children, for whom parents report that they started to be in daycare with either two years or up to 30 months belong to the early starters, while children for whom parents report that they were in daycare the first time with either three years or 31 months and later are assumed to take the more “traditional” way, which would be to be cared by the family until they are in *kindergarten*-age.<sup>33</sup> We refer to this treatment as *Early Care*. The results are, however, robust to other ways of rounding (see section 5.3). We also control for whether the parents report the month. The instrumental variable method has the advantage that it is relatively robust to measurement errors. The defined threshold is quite policy relevant, because many countries, including Germany are increasing daycare supply for children of pre-*kindergarten* age.

The focus will be on the group of children born between 1996 and 2000. These children are, consequently, five to nine years old, but children aged eight and nine are slightly under-represented (see table A14). We do not include older children, because we only have information on the regional supply of subsidized daycare for zero to three year old children, which is our exogenous variation used for identification, from 1994 on (see section 4). Some other regional information is also not available before 1996.<sup>34</sup> We focus on children that are at least five years old to make sure that the decision about (early) daycare usage was already made

---

<sup>32</sup>For children younger than three years there exists such a law since August 2013.

<sup>33</sup>It is not unusual that children start to be in *kindergarten* and as such engage the ensured slot slightly before they turn three years old.

<sup>34</sup>Most information is available from 1995 on, but not for all countries. We do not want to rely too much on imputations.

in the past.

The control-group is restricted to children who started to be in daycare in the age between 31 and 54 months. This is, firstly, because our instrument is not likely to induce any behavioral change in the childcare decision at higher starting ages. Secondly, with this strategy we ensure that each child in the sample has experienced at least some daycare at the time of observation. As shown in table 1, most of the children participating in the KiGGS have already started with daycare at latest when they turn four, anyway. As such we exclude very unusual parents, who for some unobserved reason decide absolutely against daycare. One could interpret our effects as longer compared to shorter periods of daycare, but one should keep in mind that the form of daycare for younger children differs, not only related to the ‘curriculum’, but also in terms of resources: The child-teacher ratio is for example required to be higher.

Children who live with their grandparents or in a children’s home and disabled children are also excluded, because in their life parental characteristics play a different role and the amount of non-parental care they need is very different. Children living with one parent at the time of observation remain in the sample, as long as we have the relevant information for both parents. Single parents at the time of sampling, must not have been single parents before. We control for the fact that the child lives with a separated or step parent. Given our set of control variables, this leaves us 2377 observations from which 845 belong to the early group, with slight variations in the number of observations, because of missing values for some outcomes.<sup>35</sup>

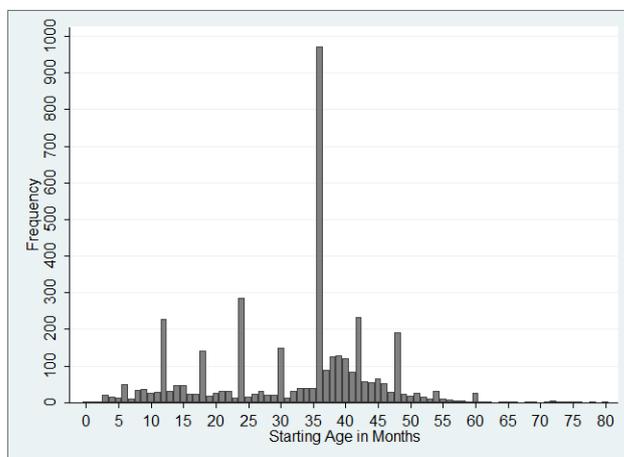


Figure 1: Distribution of the daycare starting age  
Source: KiGGS Version 4, own calculations.

The KiGGS provides several measures of physical development. We focus on indicators related to weight and fitness. The underlying measures are all drawn from an examination by a physician or a test conducted by a trained person.

Measures based on the Body-Mass-Index (BMI), calculated by dividing the weight (in kg) by the squared height (in meters), are most commonly used to determine weight problems. For children it is not appropri-

<sup>35</sup>In West Germany there are only 207 children in the treatment group and 1292 in the control group, while in East Germany there the ratio is 638 to 240.

ate to classify obesity according to the standard index for adults (BMI above 30 = obese). Therefore, obesity is defined by a BMI that lies in the 97th percentile of the relevant age-gender distribution and overweight is defined as the 90th percentile. The relevant threshold values were derived by Kromeyer-Hauschild (2005) for Germany. In our sample the probability to be overweight is on average 12 percent and the probability to be obese four percent (see table A13). The BMI variable itself is standardized to mean zero within the age-gender distribution, using half-year steps in age.<sup>36</sup> In addition we control very flexible for the age of children in months by a cubic term and height, to rule out that age-differences or growth, a rather positive development, distort our results.

Sometimes it is argued that the percentage of body fat is a better fatness indicator than the BMI, because the BMI does not make a difference between muscle, bone and fat mass (Burkhauser and Cawley 2008). However, the BMI is easier to compare and the analysis is more standardized. We provide evidence on both measures to make our results more reliable. The age and gender dependent value of the percentage of body fat is calculated according to Slaughter et al. (1988) using two skin fold measures. We use the same standardization procedure as for the BMI. In addition we calculate a risk value. Unfortunately there exist no threshold values, as it is the case for the BMI. Therefore we define the 90th-percentile in the same age-gender-group within the KiGGS to be the relevant threshold.

In addition the data contains a unique sequence of tests for fine and gross motor skills.<sup>37</sup> The gross motor skills test that can best be related to physical fitness is the side-to-side jump test. We only include the outcome of this test in the main analysis. The jumping side-to-side test was used to assess gross motor coordination under time constraint. The children were asked to perform as many jumps from side-to-side as possible during two 15 second intervals. The numbers for the two intervals were averaged (Woll et al. 2011). The test score is standardized to mean zero and standard deviation of one in the same age-gender distribution, where age is again measured in half-year steps.<sup>38</sup>

Table 1: Outcome variables

<b>Weight indicators</b>	
BMI*	$Weight(kg)/Height(m)^2$
Obese	BMI>97-percentile in the relevant age-gender group (Kromeyer-Hauschild 2005).
Overweight	BMI>90-percentile in the age-gender group (Kromeyer-Hauschild 2005).
Body Fat (BF)*	Percentage body fat (Slaughter et al. 1988).
BF-90	Body fat>90-percentile in the relevant age-gender group within KiGGS.
<b>Motor skills - test</b>	
Side Jumps (Jump)**	Number of jumps in 15 seconds, higher = better

\* Standardized to mean zero within the age-gender group (0.5 year steps).

\*\* Standardized to mean zero and standard deviation of one within the age-gender group (0.5 year steps).

<sup>36</sup>The same procedure was used by Kromeyer-Hauschild (2005) in order to derive the threshold values.

<sup>37</sup>This module is not part of the public use file. It is partly conducted by the KiGGS researches itself and partly by an additional 'Motorik-Modul' (MoMo) researcher group. The criterion for the selection of these motor tests was their capability to assess as many aspects of motor development and physical fitness as possible.

<sup>38</sup>Estimates of the effects on the other motor skills tests and socio-emotional development are also documented are provided in the appendix.

### 3.1 Estimation Strategy

The aim of this study is to provide evidence on the causal effect of early daycare usage on child health development. We can illustrate the problem in a Roy (1951) model of potential outcomes with  $Y_1$  representing the health outcome for the treated children with treatment  $D$  (*Early Care*) and  $Y_0$  the health outcome for the non-treated.<sup>39</sup>

$$\begin{aligned} Y_1 &= \alpha + \beta + u_1 & \text{if } D &= 1 \\ Y_0 &= \alpha + u_0 & \text{if } D &= 0 \end{aligned} \quad (1)$$

with  $E(u_1) = E(u_0) = 0$  in the population. In its simplest form the choice of an individual would be:

$$\begin{aligned} D &= 1 & \text{if } Y_1 - Y_0 - C &> 0 \\ D &= 0 & \text{otherwise} \end{aligned} \quad (2)$$

An individual would decide for the treatment if the outcome with the treatment is better than the outcome without the treatment plus some cost of treatment  $C$ . In this case it may actually be more appropriate to define the outcome as the parents utility from the child's health, because it is the parents, who contrast costs and benefits. We only observe one outcome per individual. However, simply taking the difference in the average outcome between the two treated and non-treated potentially provides a biased estimator of the treatment effect. Hence the parametric or non-parametric methods based on the selection on observables assumption do not identify the return to the average child attending daycare, but rather

$$\begin{aligned} E(Y|D=1) - E(Y|D=0) &= E(\beta) + E(u_1|D=1) - E(u_0|D=0) \\ &= \beta + \underbrace{E[u_1 - u_0|D=1]}_{\text{Sorting bias}} + \underbrace{E[u_0|D=1] - E[u_0|D=0]}_{\text{Selection bias}}. \end{aligned} \quad (3)$$

$ATT$

The *selection bias* refers to a situation where the outcomes of treatment and control group already differ in the untreated stage. The *sorting bias* is a bias that arises, if the gain from treatment differs between the treatment and the control group. If there is selection into treatment that depends on this gain, this induces a correlation between the treatment  $D$  and the error-term.<sup>40</sup> The most common example for these biases stems from the estimation of returns to schooling. In this case individuals in the treatment group are assumed to have higher wages even without the treatment and to be more likely to select into the treatment, because their gains from schooling are higher. One explanation is the existence of unobserved differences in abilities. For daycare the direction of this bias not clear. Children who gain more may for example stay away from

<sup>39</sup>For simplification we do not include the control variables  $X$  in this representation.

<sup>40</sup>If there is no selection bias, it is possible to identify the ATT comparing the outcomes between treatment and control group.

daycare, if parents face high unobservable costs or underestimate the treatment effect. It is also plausible that daycare is not chosen based on parental beliefs about child's gains from daycare, but for other unobserved reasons as for example working aspirations.<sup>41</sup>

Several identification issues arise when the estimation of the effect of early daycare is the objective. We are able to deal with a part of these problems using the KiGGS dataset, but may not be able to eliminate the bias completely relying on OLS or matching. The KiGGS data has the advantage that we can base our analysis on objective health measures, reported by a third, trained person, not the parent. This is important, because it has been shown that there can be errors in self-reported health (Goodman et al. 2000; Burkhauser and Cawley 2008). Measurement error in the outcome variables would not be a big problem, if it was truly random. However, in the case of the childcare decision it is not unlikely that parents who are very well informed about 'good' childcare also know better about the desired health outcomes of their children and as such overstate their health (Hills et al. 1998). We do not have this problem here.

Even though we are able to condition on a large set of covariates, we question that we observe all relevant factors. Child development indicators, in particular the pre-treatment health status, are only reported retrospectively and very approximate. We have no information about parental employment, caring and health experience and their caring attitudes. If only very healthy children select into treatment, the OLS estimator would be upward biased. If on the other hand very 'caring' or 'sporting' mothers do not select into early daycare it may be downward biased. There is some evidence that the mother's personality has an impact on childcare choice (Bjerre et al. 2011). These unobserved factors could also be related to the gains from treatment. Consequently, we cannot fully rule out reverse causality, selection and sorting issues.

Therefore, we prefer an instrumental variables estimator using the availability of subsidized daycare slots per child in the age of zero to three years in the respective region as an instrument for the starting age. In the relevant years the supply did not vary strongly over time, but there were regional differences (see section 4). This has the advantage that we are able to address the hidden part of the bias. Using this strategy, only the exogenous part of the treatment variation is used to identify the treatment effect.

The instrument  $Z$  has to fulfill the following assumptions:

1. (Conditional) Independence assumption:  $\{Y(D, Z); \forall D, Z\}, D(Z = 1), D(Z = 0)\} \perp X | Z$ : The instrument  $Z$  is as good as randomly assigned, given the the control variables  $X$ .

---

<sup>41</sup>We can also write the observed outcome as

$$Y = DY_1 + (1 - D)Y_0 = D(\alpha + \beta + u_1) + (1 - D)(\alpha + u_0) = D(\beta + u_1 - u_0) + \alpha + u_0 \quad (4)$$

$$= \alpha + \beta D + [u_0 + D(u_1 - u_0)].$$

There are three potential sources of biases. Firstly, the treatment status  $D$  could be correlated with  $u_0$ , secondly  $D$  could be correlated with  $(u_1 - u_0)$  and therefore  $\beta$  and thirdly  $\beta$  could be directly correlated with  $u_0$ . The first source arises if for example only people with genetic conditions that lead to lower weight problems select into treatment. The second source arises if families partially anticipate the treatment effect  $\beta$  and the probability to choose the treatment depends on  $\beta$ . The last mentioned bias arises if for example those with the worst outcome in the untreated state have higher returns:  $\beta$  depends on the outcome in the untreated state (see e.g. Carneiro et al. 2003).

2. Exclusion restriction  $Y(Z, D = d) = Y(Z', D = d)$  with  $Z \neq Z'$ : Conditional on the control variables there is no effect of the instrument on the outcome  $Y$  other than through the treatment.
3. Relevance of the first stage: The effect of the instrument on the treatment status is non-zero.

The supply of childcare cannot directly be influenced by the individual family. However, it could be correlated with other individual and regional factors, such as the income level. Therefore we assume that the first and the second assumption hold, but only conditional on the set of controls variable that is discussed in section 4. The relevance of the instrument is demonstrated in the same section.

This method identifies the Average Treatment Effect (ATE) only if there is no heterogeneity in the treatment effect. Equation 4 shows that if  $\beta = Y_1 - Y_0$  varies across individuals, even after conditioning on  $X$  the treatment effects vary across the population and  $Cov(\beta, D) \neq 0$ . This phenomenon is also referred to as essential heterogeneity (Heckman et al. 2006a). In this case we can only identify a local treatment effect, by imposing an additional monotonicity assumption. The LATE effects was primarily defined by (Imbens and Angrist 1994) for the case of binary instruments.

4. Monotonicity: With  $Z > Z'$ ,  $P_i(Z, X) \geq P_i(Z', X)$  for all  $i$ . The instrument moves the participation probability  $P$  for all subjects  $i$  in the same direction.<sup>42</sup>

Based on this assumption the instrumental variables estimation identifies the causal effect for those individuals who change their treatment status to *Early Care* induced by a higher subsidized daycare supply. In view of the current discussion about the increase of the provision of daycare slots, in particular in Germany, the effect of daycare usage before the age of three for those children who start to be in daycare, when the availability is higher, has a high political relevance. We assume that there are ‘compliers’, those who start using daycare early at a higher supply, but the assumption rules out that there are any ‘defiers’, here referring to parents who stop using daycare at a higher availability. All other parents are assumed to not react at all (‘always-taker’ or ‘never-taker’). Hence, we can interpret the estimated treatment effect as the causal effect for those individuals who change their treatment status to *Early Care* induced by a higher subsidized daycare supply.

The instrument in this case is continuous in order to explore the full variation in daycare supply.<sup>43</sup> The estimated effect can therefore be interpreted as a weighted average over Marginal Treatment Effects (MTE) at different levels of the instrument (Angrist and Pischke 2008). The MTE is the limit form of the LATE parameter and gives the mean effect of the treatment for those who are just indifferent between participating and not participating.<sup>44</sup> Heckman and his co-authors show how the effect estimated by linear instrumental

---

<sup>42</sup> $P_i(Z, X) \leq P_i(Z', X)$  for all  $i$  would also be possible, but not plausible in this case.

<sup>43</sup>The predictive power is as such stronger. The variation of the availability across regions in Germany is fortunately relatively high.

<sup>44</sup>The MTE is a choice-theoretic parameter that was first introduced by Björklund and Moffitt (1987).

variables is linked to the underlying MTEs (Heckman and Vytlacil 2001; Heckman et al. 2006a).<sup>45</sup>

Efficiently dealing with the binary nature of the treatment, we rely on a non-linear estimation procedure. Instead of estimating the participation equation by a linear probability model we specify a probit model. However, we then do not use the predicted probability directly in the second stage, but perform a three-step procedure, such that the predicted probability  $P(D = 1|X, Z)$  serves as the actual instrument. As such we perform a ‘stage 0’ before estimating the standard first stage and second stage. Compared to including  $P(D = 1|X, Z)$  directly in the second stage, the approach is less vulnerable against misspecification and leads to an robust estimator even if the model for  $P(D = 1|X, Z)$  is not the correct one.<sup>46</sup> We conduct the following estimation steps:

Stage 0 Estimate  $P(D = 1|X, Z)$  via probit and predict  $\hat{P}_i = \Phi(a + \mu_1 Z_i + \mu_2 X_i)$

Stage 1 Estimate  $D_i = b + \pi_1 \hat{P}_i + \pi_2 X_i + \varepsilon_i$  via OLS. Predict  $\hat{D}_i$ .

Stage 2 Estimate  $Y_i = \alpha + \beta \hat{D}_i + \gamma X_i + \varepsilon_i$ .

with  $\Phi$  being the Cumulative Distribution Function (CDF) of the standard normal distribution. This approach has the additional advantage that the standard errors of the linear Two-Stage-Least-Squares (TSLS) instrumental variables model do not have to be adjusted for the generated instrument (Wooldridge 2002).

This non-linear instrumental variable method (NTSLS) is a more efficient procedure than the standard TSLS estimator, because with the probit we use a better fitted model for the binary treatment (Newey 1990; Wooldridge 2002; Angrist and Pischke 2008). Next to Wooldridge (2002) other studies as Carneiro et al. (2011), Loken et al. (2012), Mogstad and Wiswall (2012), Fitzsimons and Vera-Hernandez (2014) and Atanasio et al. (2013) provide examples of the analysis using the predicted participation probability as the instrument. Congruently, they find substantial improvements in precision compared to the TSLS-strategy. Heckman et al. (2006a) argue that using the original instrumental directly may lead to negative weighting of some underlying MTEs, while the usage of the predicted probability as the instrument ensures that all MTEs are assigned a positive weight.<sup>47</sup> This in turn prevents from a misleading interpretation of the effect sign.

We perform subgroup analyses to narrow down the group of children for whom we identify the treatment effects via the instrumental variable estimation. Additionally to this evidence on treatment effect heterogeneity related to observable characteristic we also provide some brief evidence for the existence of effect heterogeneity related to unobservable characteristics by estimating the distribution of the MTEs. We do

<sup>45</sup>They also show that if it is possible to identify the MTE over the whole distribution, all treatment effects can be calculated by appropriately weighting the MTE. However, for these calculations a large sample and a very strong instrument that affects the participation decision over a large support are required. The linear instrumental variable weights place more weight on MTE when the cumulative distribution function of daycare usage is more strongly affected by the instrument. OLS places positive weights on every margin in the empirical support of the sample, while for linear instrumental variables the weights are zero at margins which are not affected by the instrument.

<sup>46</sup>The estimator is an efficient IV estimator if  $Var(u_0|X, Z) = Var(u_0|X)$  and probit model for D is correct.

<sup>47</sup>Using the predicted probability ensures monotonicity or uniformity.

not go further in the analysis as the sample size is too small in order to derive plausible estimates for other treatment effects based on the MTE. The patterns presented in the appendix are derived by relying on a parametric approach, assuming normality.<sup>48</sup> In this case we essentially estimate a switching regression model with a control function based on the inverse mills-ratio for the treatment and the control group (Heckman et al. 2006b; Brave and Walstrum 2014).

## 4 The Instrument and the First Stage

In Germany center-based care is highly subsidized and regulated. The provision and the fees vary between regions, but are relatively low in international comparison, because of the subsidies. The availability and not the affordability is the major problem at the daycare market (Kreyenfeld and Hank 2000).<sup>49</sup> The original functioning of these daycare centers was not to enable parents to reconcile work and family, they were rather seen as education institutions (Kreyenfeld et al. 2002). Since the 1960s the provision of center-based daycare for children aged three to six years in *kindergartens* increased strongly, such that the slot-to-child ratio was higher than 100 percent in some areas of Germany in 1998 (DJI 2008).

For children aged zero to three years the provision of daycare in particular in West Germany was very low in the relevant period, with two to three available slots per 100 children in 1998. This provision increased more strongly since the end of the 2000s, because the expansion of daycare was made an explicit political goal.<sup>50</sup> There are differences between West and East Germany. In the relevant years the supply in East Germany was much higher than in West Germany and there were many more full-day slots (Hank et al. 2004).<sup>51</sup> However, evidence suggests that there was an excess demand in West as well as in East Germany, such that we can speak of a market with strong supply side constraints (Wrohlich 2008). A higher supply of subsidized daycare for younger children can be assumed to pull the starting age to an earlier point in a child's life than the more 'traditional' age of three years.

Our main specification relies on the measure of the supply of slots in center-based care for zero- to three-year-old children at the county level ( $n \approx 470$  counties). The data is provided by the Federal Statistical Office. The actual instrument is the ratio of the available slots for zero- to three-year-old children in a county and the number of children in that age group living in that county. The slots are measured at three relevant dates: 1994, 1998 and 2002. Unfortunately, there is no statistic on yearly basis, such that we have to (linearly) impute between these years. The results are, however, robust to these imputations, as there are

---

<sup>48</sup>The sample is also too small for a reasonable non-parametric analyzes.

<sup>49</sup>The decision about the provision is most often made by municipalities. The funding is also provided by local authorities. The staff is trained and there exist relatively high quality standards compared to other countries (see Kreyenfeld and Hank (2000) for an international comparison and Blau (2001) on the quality problem in particular in the US.), but the slots are limited and opening hours are not very flexible.

<sup>50</sup>However, still less than 15 percent of slots are full-time care.

<sup>51</sup>In some areas of West Germany the provision was around three percent, in East Germany at a level of 36 percent. In the following years this number increased only slightly to 6.8 percent for West Germany and 36.7 in East Germany (Destatis 2007). The provision of subsidized care even decreased after the reunification (Kreyenfeld et al. 2002).

few changes over time in this period (Hank et al. 2004).

This instrument would not be valid if the level of provision was correlated with the child's health. A direct connection is rather implausible, but there may be indirect channels through which a correlation arises, such as the regional wealth level, demand or norms, because the decision about the provision of subsidized daycare was mainly taken by regional authorities. In order to meet assumptions one and two, we control for income (per capita), the unemployment rate, the fraction of the working population, the fraction of working women aged 30 to 50 years, the fraction of foreigners, the number of children at the age of zero to three years and the population density.<sup>52</sup> The KiGGS includes an additional measure of urbanization. We include state dummies, because the state government decides about the financial support for childcare. Standard error are clustered at the county level. Given the different caring systems and related norms in East and West Germany we include interactions between the regional characteristics and an East Germany dummy. Hence, we allow for example for different effects of regional income on daycare supply. In addition we control for a set of individual and family specific characteristics. These are a set of basic child characteristic, child early health indicators, parental socio-economic background variables, parental health and health behavior proxies, measures of family composition and indicators of the caring mode (see table A1). Including these variables we argue that we block the potentially channels through which the instrument could be correlated with the outcome.

Figure 2 pictures the distribution of slots for zero- to three-year-old children in the year 1998. It confirms that the level between former East and West Germany differs remarkably. However, as we include state fixed effects, we only use the variation within a state as the source of exogenous variation for early daycare usage. Table 2 gives the distribution of the slot-child-ratio for zero- to three-year-old children within states in the KiGGS data. The KiGGS is sampled at 167 points such that we are able to use a large fraction of the regional variation.

A few studies successfully use a similar instrument in the context of caring decisions. For Germany there are the studies of Felfe and Lalive (2012; 2013), analyzing the effect of early daycare usage on child cognitive and behavioral development. There is also the study of Kröll and Borck (2013), who use the regional variation to exploit the effect of daycare on maternal well-being and mother-child interactions. Dustmann et al. (2012) estimate the effect of daycare in *kindergarten*-age on child development. Datta Gupta and Simonsen (2010; 2012) perform a similar instrumental variables analysis estimating the effect of different types of care in Denmark.

The instrument in the mentioned German studies is, however, not exactly the same as ours. Felfe and Lalive (2012) and Kröll and Borck (2013) use the ratio of children in center-based care divided by the children in the relevant age group living in the county, while we use the actual availability of slots in center-based care

---

<sup>52</sup>These indicators were measured in year of the child's birth. Most county-level characteristics change only slowly over time.

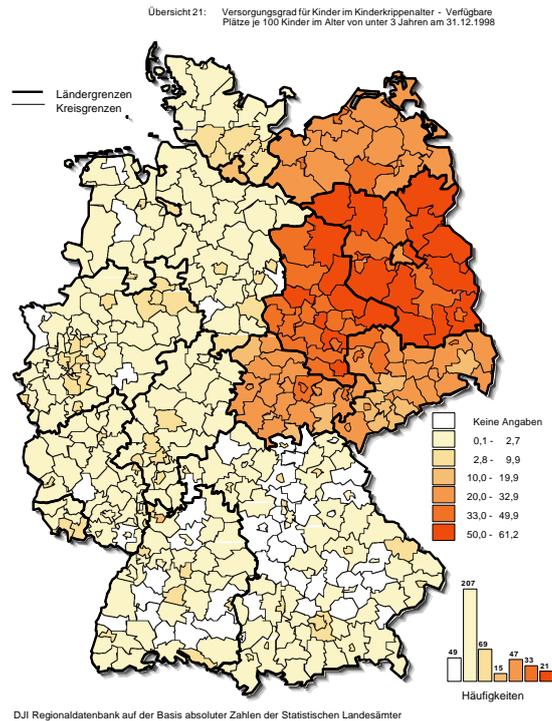


Figure 2: Distribution of slot to child ratios children aged 0-3 years across German regions (1998)  
Source: Zahlenspiegel, Deutsches Jugendinstitut, 2002.

divided by the number of children. Our instrument, consequently, actually measures supply.<sup>53</sup>

Dustmann et al. (2012) and Felfe and Lalive (2013) do not combine their data on individual observations with administrative data, but use the average in daycare usage of all children in their sample at one child's place of living as a proxy for availability.<sup>54</sup> We exploit this approach as a robustness check. We average the usage of *Early Care* of all four- to 13-year-old children at the sample point, excluding the individual's own observation. As daycare supply is rationed, demand can be assumed to equal supply and as such we get a second measure of availability. This measure does not only include subsidized center-based care, but also other care alternatives that are not part of the official statistics.

A last worry could be the selection of parents into areas of high childcare supply. The KiGGS data unfortunately does not provide data on mobility. Based on administrative data we can only provide crude evidence for the rejection of this hypothesis.<sup>55</sup> We prefer to refer to Felfe and Lalive (2012; 2013), who are able to perform this exercise based on individual data. They argue that only a few parents move and that mobility

<sup>53</sup>Until 2002 the Federal Statistical Office provided information on the supply of slots in daycare centers, from 2006 on they provide the number of children who are in childcare. Even if the market is very restricted, this does not have to be the same.

<sup>54</sup>Both also capture a variation generated by the expansion over time.

<sup>55</sup>Conditional on our set of controls variables individuals in the relevant age are not more likely to move into areas of higher daycare supply. The Federal Statistical Office only provides the moving information for broad age groups. Aged 30 to 50 years should include most of the parents. We use a binary indicator which equals one in a county in which more people move in than people move out in a given year. The value zero refers to the opposite case. Results available on request.

Table 2: Distribution of the child-slot ratio within states

State (' <i>Bundesland</i> ')	Mean	Min.	Max.	Sd.
<i>West</i>				
Schleswig-Holstein	0.019	0.007	0.033	0.010
Hamburg	0.119	0.117	0.124	0.002
Lower Saxony	0.020	0.000	0.060	0.020
Bremen	0.077	0.075	0.084	0.003
North Rhine-West	0.021	0.000	0.062	0.015
Hesse	0.025	0.006	0.078	0.020
Rhineland-Palatinate	0.023	0.002	0.073	0.024
Baden-Wuerttemberg	0.014	0.000	0.084	0.019
Bavaria	0.011	0.000	0.075	0.018
Saarland	0.029	0.007	0.043	0.012
Berlin	0.319	0.307	0.349	0.010
Brandenburg	0.509	0.409	0.624	0.057
Mecklenburg-West Pomerania	0.330	0.217	0.511	0.080
Saxony	0.259	0.117	0.401	0.062
Saxony-Anhalt	0.501	0.429	0.602	0.047
Thuringia	0.280	0.175	0.433	0.054
Total	0.149	0.000	0.624	0.183

Source: KiGGS Version 4 & Federal Statistical Office, own calculations.

decisions are unrelated to the number of daycare slots in a German county. Felfe and Lalive (2012) also find that daycare supply is conditionally orthogonal to a range of key determinants of child development.<sup>56</sup> Their birth cohorts are 2002 to 2008. In this time period childcare was already a rather big topic in Germany. Their arguments are therefore even more likely to hold for an earlier time period.

#### 4.1 The First Stage

Table 3 reveals that the instrument has a strong impact on the participation decision. This is the case using the slot-child ratio directly and using the estimated propensity score. The effects are highly significant in both cases. In section 5 we provide for each analyzed sample also the F-statistics which detects by a rule of thumb a weak instrument problem at a value of 10 or lower (Staiger and Stock 1997).<sup>57</sup> For the predicted participation probability this number outgrows the critical value by far. This instrument has a much stronger predictive power than the original variable. This should lead to more precisely estimated effects in the second stage. For this reason Mogstad and Wiswall (2012) call the predicted probability the “*efficient instrument*”.<sup>58</sup>

The F-statistics of above 20 in the full sample, using the slot-child ratio directly (TSLS), reassures that there is no weak instrument problem, even for the standard linear model. Hence we can assume that the

<sup>56</sup>They mention birth outcomes, such as birth height, proxies for parenting style and maternal preferences.

<sup>57</sup>The information is always based on the estimation for BMI. In addition we provide the value of the Partial R-Squared, a measure that provides information about instrument relevance by netting out the effect of the other variables and as such isolating the correlation between the endogenous regressor and the excluded instrument. What matters for instrumental-variables estimation is whether the component of *Early Care* that is orthogonal to the other regressors can be explained by the component of the predicted value of *Early Care* that is orthogonal to the predicted values of the other regressors in the model. Similar to Shea (1997)‘s Partial R-squared.

<sup>58</sup>They find strong efficiency gains of the non-linear model in particular for low probability events.

non-linearities are not the sole mechanism driving identification in the NTSLs estimation.<sup>59</sup> The instrument performs weaker in some of the analyzed subsamples.

Table 3: First Stage TSLS and NTSLs of *Early Care*

	All	West	East	Girls	Boy	BMI father > 25	BMI father ≥ 25	Household Income High	Household Income Low
Slot-Child Ratio	1.071*** (0.314)	1.959** (0.829)	0.954*** (0.325)	1.137*** (0.343)	0.893* (0.471)	0.650* (0.332)	1.230*** (0.374)	0.471 (0.541)	1.257*** (0.369)
Predicted Probability	1.220*** (0.084)	1.278*** (0.154)	1.212*** (0.221)	1.160*** (0.116)	1.078*** (0.111)	1.067*** (0.091)	1.261*** (0.107)	1.148*** (0.117)	1.118*** (0.099)
Observations	2377	1487	878	1202	1158	999	1347	753	1599

*Notes:* Estimates are based on OLS estimations controlling for the set of variables defined in 'Child characteristic', 'Parents and Family' and 'Regional controls' as described in table A1. Inference is based on standard errors that allow for clustering at the school district level.\*/ \*\*/ \*\*\*: Significance at the 10-/5-/1-percent level.

*Source:* KiGGS Version 4 & Federal Statistical Office, own calculations.

Further coefficients of the first stage reveal some selection patterns (table 3). The effects do not all point to a coherent story. High educated mothers and mothers who are still in training seem to be more likely to use daycare early. The opposite is the case for fathers with vocational training, in particular if they have *Abitur*. Small families and single or separated parents use daycare earlier. The same is the case for mothers who breastfed, but also for parents who smoke more regularly at home. The unemployment rate at the family's place of living has only a significantly positive associated with *Early Care* usage in East Germany. A place with higher population density daycare is use earlier. We also find as positive impact of whether the parent reported the starting months. It seems that the early users are better informed about the exact starting age. All other regional effects are small and not statistically different from zero. Individual attendance is hence unrelated to most county level variables once we account for daycare supply. If we assume that a bias due to observables is a good proxy for the bias due to unobservables, an argument brought forward by Altonji et al. (2005), we could conclude that a bias due to unobserved county variables is unlikely.

## 5 Results

### 5.1 Treatment Effects - Early Universal Daycare

Table 4 gives the estimation results for the full sample. The unconditional OLS coefficients displayed in the last row indicate a tendency of worse physical development if children are in daycare earlier. However, controlling for available observable characteristics turns these tendencies around. The estimates are not statistically different from zero, but the change in sign indicates that there are observables differences between the treatment and control group, such that we have to assume that there is also selection on unobservables. The NTSLs estimates reveals are very different picture for the children 'at the margin'. There is a reduction in the level of the BMI and the percentage of body fat, by more than one point and more than two percent.

<sup>59</sup>(See Carneiro et al. 2003) for a similar argumentation.

The probability to be overweight is 13 percentage points lower and the probability to be obese 9 percentage points lower. In a similar range we find a reduction in the probability to belong to the 90th-percentile of the body-fat-distribution by 12 percentage points. This picture of a better physical development is confirmed by a positive effect of *Early Care* on the number of side-to-side jumps. Within the relevant age-gender group the children in *Early Care* score 49.6 percent of a standard deviation better on average. These are effects of a size that lies between not having an overweight and not having an obese mother. It is a well-known that parental weight problems a crucial determinant for child's weight problems (see e.g. Danielzik et al. 2004).

Table 4: Effect of *Early Care* - All children aged 5-9

	Body-Mass-Index			Body Fat		Motor Jump
	BMI	Overweight	Obese	BF	BF>p90	
<i>NTSLS - Main instrument</i>						
<i>Early Care</i>	-1.064** (0.443)	-0.130** (0.065)	-0.094** (0.042)	-2.404** (1.105)	-0.116* (0.061)	0.467** (0.219)
<i>TSLs - Main instrument</i>						
<i>Early Care</i>	-1.176 (1.265)	-0.158 (0.203)	0.080 (0.139)	-1.833 (3.402)	-0.103 (0.192)	1.145* (0.664)
<i>NTSLS - 2nd instrument</i>						
<i>Early Care</i>	-0.936** (0.441)	-0.135** (0.066)	-0.103** (0.041)	-1.951* (1.123)	-0.113* (0.062)	0.262 (0.227)
<i>TSLs - 2nd instrument</i>						
<i>Early Care</i>	-0.588 (1.104)	-0.246 (0.184)	0.030 (0.110)	-0.770 (3.011)	-0.130 (0.173)	-0.047 (0.552)
<i>OLS</i>						
<i>Early Care</i>	-0.066 (0.130)	-0.012 (0.021)	-0.006 (0.013)	-0.354 (0.335)	-0.005 (0.019)	0.040 (0.056)
<i>OLS - no controls</i>						
<i>Early Care</i>	0.114 (0.104)	0.003 (0.015)	-0.004 (0.009)	0.304 (0.289)	0.006 (0.013)	-0.034 (0.058)
Observations	2372	2372	2372	2348	2348	2335

F-stat.(NTSLS): 202.09 , F-stat.(TSLs): 21.64, Part. R-sq.(TSLs): 0.009

F-stat.(NTSLS 2): 191.61, F-stat.(TSLs 2): 29.27, Part. R-sq.(TSLs 2): 0.013

*Notes:* All estimates are based on IV estimations controlling for the set of variables defined in 'Child characteristic', 'Parents and Family' and 'Regional controls' as described in table A1.

Inference is based on standard errors that allow for clustering at the county level.

\* / \*\* / \*\*\*: Significance at the 10-/5-/1-percent level.

*Source:* KiGGS Version 4 & Federal Statistical Office, own calculations.

The estimated effects using the second instruments are very similar and confirm all effects, but the one on motor skills development. The point estimates are very similar in all cases. The standard TSLs estimates for both instruments are also in line with the point estimates, but the effects are as expected much less precisely estimated. Only in the case of obesity the point estimates hint into another direction, but the estimated effects are neither close to be significantly different from zero nor significantly different from the NTSLS estimates. They also contradict the tendencies found for the other outcomes via TSLs.

It is not an unusual finding that OLS estimates are smaller than instrumental variables estimates, at least

in the estimation of returns to education.<sup>60</sup> It is also not the first case in which OLS findings for early care are much weaker than the instrumental variables estimates.<sup>61</sup> One interpretation is that with the instrumental variable method we identify a treatment effect for the marginal individual, while OLS identifies a - potentially biased - estimate of the average treatment effect.

If children whose parents are just indifferent between using or not using daycare early, gain more than the children in daycare gain on average this may explain the gap. A potential reasoning would be that individuals who face a relatively high cost of *Early Care*, gain most.<sup>62</sup> As costs, including search costs, related to daycare for young children are relatively high compared to daycare costs for older children in most areas in Germany, this may be a plausible chain of argumentation here. In particular, if we keep in mind that in the observation period parental leave payments last for two years, but only if one parent did not work more than 19 hours (or 30 since 2001) and the household income did not exceed a certain threshold. Consequently, it is quite plausible that incentives to rely on daycare early and thus the costs of using *Early Care* varied between families.<sup>63</sup>

It should be kept in mind that this is a weighted effect, averaging over the MTEs for different groups. Therefore, we cannot infer anything about the OLS bias, comparing the estimated OLS effects with the effects derived by the instrumental variables strategy.

As these estimates cannot be interpreted as ATEs, but are effects for specific subgroups, we next try to figure out for which groups these positive treatment effects exist. However, it turns out that we can not conduct reliable estimates of causal effects for all subgroups, because only some of these groups actually react to the differences in subsidized center-based care supply. We show in the following that the reaction is much stronger in families without high household income and with an overweight father. This finding prevents us from discussing effect heterogeneity in its usual way, by comparing effect sizes between groups, but it allows us to better locate the population for which the estimated effect is relevant.

## 5.2 For Whom Are We Estimating the Effects?

Table 5 gives the local treatment effects estimated separately for boys and girls. The first stage estimates presented in table 3, already suggest that the strength of the instrument weakens if we split the sample by gender. This is certainly the case, because the variation decreases if fewer observations are included. However, based on this rather weak instrument, which at least for girls would still pass the 'rule of thumb' test for weak instruments, we would conclude that both boys and girls gain from *Early Care* by reduced weight problems and also a better motor skills performance. For boys the second instrument performs better, according to the F-statistics, but the effects are very similar (not shown). The point estimates for girls

---

<sup>60</sup>See Card (1999) for an overview.

<sup>61</sup>See for example Felfe and Lalive (2012; 2013).

<sup>62</sup>This could be the case, if there were credit constraints (Card 2001).

<sup>63</sup>Related to the returns to education Carneiro and Heckman (2002) discuss comparative advantage and other channels that could explain this pattern.

and boys diverge slightly, but do not differ statistically.

Table 5: Effect of *Early Care* - Subgroup analysis: Gender

	Body-Mass-Index			Body Fat		Motor
	BMI	Overweight	Obese	BF	BF>p90	Jump
<i>NTSLS-Girls</i>						
<i>Early Care</i>	-1.139** (0.572)	-0.226*** (0.084)	-0.133** (0.058)	-2.074 (1.494)	-0.137* (0.081)	0.388 (0.299)
Observations	1200	1200	1200	1187	1187	1185
F-stat.(NTSLS): 106.08 , F-stat.(TSLs): 12.41, Part. R-sq.(TSLs): 0.011						
<i>NTSLS-Boys</i>						
<i>Early Care</i>	-1.534** (0.680)	-0.104 (0.097)	-0.043 (0.064)	-2.631 (1.632)	-0.153* (0.092)	0.320 (0.290)
Observations	1172	1172	1172	1161	1161	1150
F-stat.(NTSLS): 101.63 , F-stat.(TSLs): 6.99 , Part. R-sq.(TSLs): 0.006						

*Notes:* All estimates are based on IV estimations controlling for the set of variables defined in ‘Child characteristic’, ‘Parents and Family’ and ‘Regional controls’ as described in table A1. Inference is based on standard errors that allow for clustering at the county level.

\* / \*\* / \*\*\*: Significance at the 10-/5-/1-percent level.

*Source:* KiGGS Version 4 & Federal Statistical Office, own calculations.

The finding on the strength of the instrument is even a little worse if we split the sample into East and West Germany. This may not be surprising, because we include fewer regions per subsample. The difference in the F-statistics between the NTSLS and TSLs are much smaller in East than in West Germany. The improvement in precision by using the NTSLS may thus be much larger in West Germany, where *Early Care* is less common. This result is coinciding with the insights derived by Mogstad and Wiswall (2012), who argue that the increase in efficiency by NTSLS is largest in the case of low probability events. However, the effects in both areas of Germany are not of high significance and not statistically distinguishable. In West Germany the second instrument performs better and the first instrument also performs better if regions with zero supply are excluded, but the results do not differ substantially. Even if it is of interest to go further into the difference between these two regions, it is obvious that the estimators are too imprecise for a reasonable comparison. The body-fat measures point to larger effects in West Germany, but all other point estimates are of similar size. At least, none of these estimators indicates that one or the other group suffers from *Early Care*.

The effect on individuals, who react most strongly to the change in the instrument are implicitly weighted more heavily. Hence we are also interested in whether we can infer if we estimate an effect for disadvantaged or advantaged children. We define a wealthy group, which are those with a monthly household income of at least €3000 and a less wealthy group with a lower household income. A look at the first stage regression for these two groups already reveals that the instrument has a very weak effect in the group of wealthy families (table 3). The instrument has a much stronger impact for less wealthy families. Excluding the children of wealthy families does actually not change the estimated effects compared to the full sample. For most outcomes the effect size is even larger (table 7). This finding is absolutely in line with the pattern provided

Table 6: Effect of *Early Care* - Subgroup analysis: Region

	Body-Mass-Index			Body Fat		Motor
	BMI	Overweight	Obese	BF	BF>p90	Jump
<i>NTSLS-West</i>						
<i>Early Care</i>	-0.750 (0.592)	-0.053 (0.096)	-0.015 (0.058)	-3.682** (1.739)	-0.182* (0.100)	0.079 (0.334)
Observations	1494	1494	1494	1474	1474	1467
F-stat.(NTSLS): 126.84 , F-stat.(TSLs): 7.51 , Part. R-sq.(TSLs): 0.005						
<i>NTSLS-East</i>						
<i>Early Care</i>	-0.466 (0.894)	-0.099 (0.141)	0.007 (0.095)	-1.000 (2.344)	-0.073 (0.132)	0.887** (0.413)
Observations	878	878	878	874	874	868
F-stat.(NTSLS): 35.34, F-stat.(TSLs): 10.92, Part. R-sq.(TSLs): 0.013						

*Notes:* All estimates are based on IV estimations controlling for the set of variables defined in 'Child characteristic', 'Parents and Family' and 'Regional controls' as described in table A1. Inference is based on standard errors that allow for clustering at the county level.

\* / \*\* / \*\*\*: Significance at the 10-/5-/1-percent level.

*Source:* KiGGS Version 4 & Federal Statistical Office, own calculations.

by the subgroups analysis via OLS. These OLS estimates point into the direction that *Early Care* leads to fewer weight problems and better fitness for children in middle and lower income households (see table A2). Based on the instrumental variable strategy we can confirm this finding. Additional results not presented here reveal a similar pattern for maternal education.

Table 7: Effect of *Early Care* - Subgroup analysis: Household Income

	Body-Mass-Index			Body Fat		Motor
	BMI	Overweight	Obese	BF	BF>p90	Jump
<i>Monthly household income ≥ €3000</i>						
<i>Early Care</i>	0.808 (0.577)	0.162* (0.095)	0.034 (0.055)	0.122 (1.713)	0.015 (0.084)	0.628* (0.349)
Observations	754	754	754	747	747	745
F-stat.(NTSLS): 69.51, F-stat.(TSLs): 0.80, Part. R-sq.(TSLs): 0.001						
<i>Monthly household income &lt; €3000</i>						
<i>Early Care</i>	-1.268** (0.573)	-0.189** (0.089)	-0.099* (0.058)	-1.875 (1.364)	-0.115* (0.069)	0.374 (0.258)
Observations	1618	1618	1618	1601	1601	1590
F-stat.(NTSLS): 148.37, F-stat.(TSLs): 24.48 , Part. R-sq.(TSLs): 0.016						

*Notes:* All estimates are based on IV estimations controlling for the set of variables defined in 'Child characteristic', 'Parents and Family' and 'Regional controls' as described in table A1. Inference is based on standard errors that allow for clustering at the county level.

\* / \*\* / \*\*\*: Significance at the 10-/5-/1-percent level.

*Source:* KiGGS Version 4 & Federal Statistical Office, own calculations.

Related to health development it is further of interest whether we estimate an effect for children who are at risk from a health point of view. Interestingly, we do not find large differences related to the mother's BMI, but a strong pattern related to the father's BMI. The results in table 8 show that we only find treatment effects that are statistically different from zero in the group of children with a father who is overweight, defined as having a BMI that is larger than 25. The estimated effects are even larger than in the combined

sample. A comparison of the F-statistics in the two subsamples shows that the explanatory power of the instrument is also much stronger in this group. Families with high BMI fathers seem to react much stronger to differences in subsidized center-based care supply. Our estimates should, consequently, be interpreted as effects that are strongly driven by this subgroup.

Table 8: Effect of *Early Care* - Subgroup analysis: Paternal BMI

	Body-Mass-Index			Body Fat		Motor Jump
	BMI	Overweight	Obese	BF	BF>p90	
<i>BMI father ≤ 25</i>						
<i>Early Care</i>	-0.269 (0.444)	-0.022 (0.064)	-0.013 (0.028)	-0.442 (1.262)	-0.010 (0.069)	0.441 (0.290)
Observations	1018	1018	1018	1006	1006	1002
F-stat.(NTSLS): 105.40, F-stat.(TSLs): 2.97, Part. R-sq.(TSLs): 0.003						
<i>BMI father &gt; 25</i>						
<i>Early Care</i>	-1.294** (0.648)	-0.101 (0.096)	-0.150** (0.070)	-2.715* (1.536)	-0.184** (0.089)	0.404 (0.260)
Observations	1354	1354	1354	1342	1342	1333
F-stat.(NTSLS): 125.26, F-stat.(TSLs): 16.75, Part. R-sq.(TSLs): 0.013						

*Notes:* All estimates are based on IV estimations controlling for the set of variables defined in ‘Child characteristic’, ‘Parents and Family’ and ‘Regional controls’ as described in table A1. Inference is based on standard errors that allow for clustering at the county level.

\* / \*\* / \*\*\*: Significance at the 10-/5-/1-percent level.

*Source:* KiGGS Version 4 & Federal Statistical Office, own calculations.

These findings are not implausible. Parents with higher socio-economic background, which are on average also those with fewer weight problems, may be more likely to belong to the ‘always-takers’ or ‘never-takers’. Their decision may be less dependent on subsidized daycare supply, because at a higher income level families may have better access to private solution. We should keep in mind that we are not able to provide reliable causal evidence for families with high income and families with a father with normal BMI, because the instrument performs weak in these groups.

The estimation of the distribution of the MTE, pictured in section H, reveals that it is likely that there is also heterogeneity in the treatment effects related to unobservables. It seems for example that the reduced BMI effect arises in areas of high propensity score or for families who just decide to participate at high unobservable cost ( $u_D$ ). The estimated effects should therefore definitely be interpreted as local treatment effects. However, the derived confidence bands also reveal that we are not able to estimate a distribution that is significantly different from zero, even under the strong assumption of normality. We have to leave the exercise of deriving additional treatment parameters based on distribution of the MTEs to future research based on a larger sample.

### 5.3 Sensitivity Analysis

As the results rely on an instrument that is partly a linear imputation between three points in time, we check the robustness of the results to different specifications of the instrument. Firstly, we only use the

slot-child ratio in 1998. Secondly, because of the differences between East and West Germany we include the interaction between the East Germany dummy and the instrument as an additional instrument. Thirdly, we include a squared term of the slot-child ratio in addition to the linear term (see table E). We also specified the stage zero as a logit model. The results are very robust to these changes.

We also checked the robustness to different specifications. With a limited set of controls variables none of the results are contradicting the main findings, but more control variables increase as expected the precision of our estimates.<sup>64</sup> As there were worries that the provision of subsidized care is correlated with weight problems via some unobservables factors, we include additional regional variables, in particular the mean BMI of very young children (aged zero to two) and the mothers' and fathers' average BMI at the sampling point (see table A9). The results are robust to the inclusion of these measures. We take this as additional evidence in favor of the assumption that our estimates are not driven by regional difference in weight problems.

We also provide comparison results for socio-emotional development, measured by the Strengths and Difficulties Questionnaire (SDQ) (see table A11) and on the other motor tests in the KiGGS. For the well known, but parental reported measures of SDQ, we only find a significant reduction of conduct problems. Thus results are in line with the research that focuses on the effect of universal care on these developmental outcomes, such as Datta Gupta and Simonsen (2012) or Felfe and Lalive (2012; 2013). For the additional motor tests, that also examine fine motor skills, we find even some negative, but no significant, tendencies. Consequently, our estimates do not seem to capture unobserved factors that make children better off in all dimensions. An additional estimate ensures as that the children are not more likely to be underweight (belong to the lowest 10 percent of the weight distribution).

Even though the chosen cutoff point of 30 months is well justified, there may be worries about the rounding procedure. Therefore, we estimate the effects again using a three year and a traditional 'mathematical' rounding threshold. The results given in table A8 reveal that there are only slight changes in the point estimates and significance, but the results of baseline specification are confirmed.

Reducing the sample to children of primary-school-age (age six to nine years), for whom the decision about and the use of daycare lies in the past, shows that the effects go through within this sample (table E). We have to keep in mind that these effects do not allow to draw conclusions about the persistence of effects, because the individuals reacting to differences in daycare supply could actually be the same in the two samples.

We also checked the sensitivity of our results to the choice of the subsample. We included children, for whom the parents report that the child was only cared by a childminder, children who were born in 1995<sup>65</sup> and children who started to be in daycare, when they were 54 months old or older, thus relatively late. In all

---

<sup>64</sup>The results are also robust to excluding small city states with low variation and the small group of nine-year old children (not shown). If we include district dummies (regional authorities between counties and states) we still find similar effects, but the estimates are nearly not significantly different from zero, because of the strongly reduced regional variation.

<sup>65</sup>Imputing missing regional values by following years.

three case the point estimates are very similar to the main results, but the estimates are a little less precise. This may be a result of a weaker instrument. The availability of subsidized center-based care for children aged zero to three can be expected to have a weaker effect on the decision to make use of a childminder and in the group that for some unobservable reason belongs to the late starters. The value of the partial R-squared points in this direction. In the earlier years there were more areas in West Germany with zero supply, such that the distribution in the instrument is more strongly truncated.

## 6 Discussion and Conclusion

In this paper we estimate effect of universal daycare usage in pre-*kindergarten*-age, defined by a starting age with 30 months or younger, on child physical development using a representative health dataset for children in Germany. Bad physical development is defined as weight and fitness problems. As recommended by the literature we do not only base our judgment of weight problems on the Body-Mass-Index, but also use indicators based on the percentage of body fat (Burkhauser and Cawley 2008). We include a test of gross motor skills as a proxy for the child's fitness and as such provide a broad picture of the child's physical development. All health indicators are recorded by a trained person and as such not prone to a bias due to parental reporting. We compare children who are in daycare early with those who start later, such that the estimates could also be interpreted as the effect of the duration in daycare before school enrollment.

We use an instrumental variable strategy to deal with the selection into early daycare. We exploit regional differences in the supply of subsidized center-based care for zero- to three-year-old children as the source of exogenous variation in early daycare usage. Our focus is on short- to mid-term effects, such that we estimate effects for five- to nine-year-old children.

After conditioning on a large set of early health, parent, family and regional characteristics, we find no significant developmental difference between children that were in daycare earlier and those who started later. However, all effects point to a reduction of weight problems and a better performance in the gross motor test. Splitting the sample into income groups we even find a significant lower occurrence of weight problems in the middle and lower income groups.

The estimated weighted local average treatment effects are derived by a non-linear instrumental variable strategy, for which we provide additional evidence that it is more efficient in case of a binary treatment than the standard linear instrumental variable strategy. The effects are much more precisely estimated, but the point estimates coincide. For children of parents who are just indifferent between using daycare earlier and later, but chose to enroll their child in daycare early induced by a relatively higher supply of subsidized center-based care, we find a significant reduction of weight problems and also a significant better performance in the test for gross motor skills.

We are not able to provide very reliable statistical evidence for difference by gender or region, because the

instrument becomes relatively weak and estimates are rather imprecise if we split the sample. The estimated effects do not hint to differences by gender or between East and West Germany. As the instrument does not change behavior for all individuals, we try to identify the groups that are most strongly affected. We find that the decision to use daycare early is very strongly influenced by differences in subsidized center-based care supply in families that do not belong to the highest income group and families with an overweight father. In these groups the treatment effects are particularly strong.

Our results contradict Herbst and Tekin (2011) who find negative effects of maternal employment accompanied by center-based care usage on weight problems in the US. Herbst and Tekin (2011), however, analyze the effect of a daycare subsidy for disadvantaged single mothers, which strongly increased labor supply and may also have led to the use of low quality daycare. The German daycare system is said to be of relatively high quality and was not primarily set up in order to increase female working hours. Therefore, many slots are part-time, in particular in end of the 1990s and early 2000s. There is evidence suggesting that shorter hours and higher quality should be assumed to make a crucial difference.

As most of the children in our sample were in center-based care, our findings coincide with the insights given by Datta Gupta and Simonsen (2012). They show that center-based care in Denmark is at least as good as home care related to socio-emotional development. Felfe and Lalive (2012) and Felfe and Lalive (2013) in a later time period also find strong positive effects of early daycare usage in Germany, in particular on the school readiness tests. Hence our positive results fit well in the existing literature.

We also agree with the findings for the early education programs in the US. The estimated effects are somewhat smaller than those of Frisvold (2006), Carneiro and Ginja (2008) or Campbell et al. (2014). This is not surprising as these programs included some guidelines that aim to change the children's diet as well as their exercising behavior and were targeted to families with low socio-economic background, which are more likely to have weight problems. In the present paper we analyze a universal type of daycare, such that the group of children is more mixed. Still, it could be the case that our instrumental variable strategy also captures children that are at risk. Education in terms of eating and exercising behavior is, however, not an explicit object in German daycare centers.

Due to data restrictions, we cannot infer much about the factors driving our findings. The question whether differences in the diet and as such a persistence change in eating habits (see e.g. Birch 1999), a change in daily exercise by being more active with other children or even other changes in behavior and preferences are responsible for the effects has to be left to future research, as our dataset does neither provide information about the daycare intensity nor about the daily routine in the daycare centers.

Despite these limitations, an appropriate and justifiable interpretation of our findings would be that we detect additional evidence for the hypothesis that early daycare usage does not harm children, but may even foster the development, at least of some children, in all kind of areas, including a reduction of weight and fitness problems.

## References

- ALTONJI, J. G., T. E. ELDER, AND C. R. TABER (2005): "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." *Journal of Political Economy*, 113, 151.
- ANDERSON, P. M., K. F. BUTCHER, AND P. B. LEVINE (2003): "Maternal Employment and Overweight Children," *Journal of Health Economics*, 22, 477–504.
- ANDREWS, R. J., P. JARGOWSKY, AND K. KUHNE (2012): "The Effects of Texas's Targeted Pre-Kindergarten Program on Academic Performance," *NBER Working Paper No. 18598*.
- ANGRIST, J. D. AND J.-S. PISCHKE (2008): *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press.
- ATTANASIO, O. P., V. D. MARO, AND M. VERA-HERNÁNDEZ (2013): "Community Nurseries and the Nutritional Status of Poor Children. Evidence from Colombia," *The Economic Journal*, 123, 1025–1058.
- BAKER, M., J. GRUBER, AND K. MILLIGAN (2008): "Universal Child Care, Maternal Labor Supply, and Family Well-Being," *Journal of Political Economy*, 116, 709–745.
- BARNETT, L. M., P. J. MORGAN, E. VAN BEURDEN, AND J. R. BEARD (2008): "Perceived Sports Competence Mediates the Relationship between Childhood Motor Skill Proficiency and Adolescent Physical Activity and Fitness: A Longitudinal Assessment," *International Journal of Behavioral Nutrition and Physical Activity*, 5, 40.
- BÄUERLEIN, K., C. LINKERT, E. STUMPF, AND W. SCHNEIDER (2013): "Kurz- und langfristige Effekte außerfamiliärer Kleinkindbetreuung auf die kognitive und sprachliche Entwicklung unter besonderer Berücksichtigung der Betreuungsqualität," *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 45, 57–65.
- BAUM II, C. L. (2003): "Does Early Maternal Employment Harm Child Development? An Analysis of the Potential Benefits of Leave Taking," *Journal of Labor Economics*, 21, 409–448.
- BECKER, G. S. (1965): "A Theory of the Allocation of Time," *The Economic Journal*, 75, 493–517.
- BENSON, L. AND M. MOKHTARI (2011): "Parental Employment, Shared Parent-Child Activities and Childhood Obesity," *Journal of Family and Economic Issues*, 32, 233–244.
- BERGER, E. M. (2013): "Happy Working Mothers? Investigating the Effect of Maternal Employment on Life Satisfaction," *Economica*, 80, 23–43.
- BERGER, L. M., J. HILL, AND J. WALDFOGEL (2005): "Maternity Leave, Early Maternal Employment and Child Health and Development in the US," *The Economic Journal*, 115, F29–F47.

- BERLINSKI, S., S. GALIANI, AND P. GERTLER (2009): “The Effect of Pre-Primary Education on Primary School Performance,” *Journal of Public Economics*, 93, 219–234.
- BERNAL, R. AND M. P. KEANE (2011): “Child Care Choices and Children’s Cognitive Achievement: The Case of Single Mothers,” *Journal of Labor Economics*, 29, 459–512.
- BIRCH, L. L. (1999): “Development of Food Preferences,” *Annual Review of Nutrition*, 19, 41–62.
- BJERRE, L., F. H. PETER, AND C. K. SPIESS (2011): “Child Care Choices in Western Germany also Correlated with Mother’s Personality,” *DIW Economic Bulletin*, 1, 20–26.
- BJÖRKLUND, A. AND R. MOFFITT (1987): “Estimation of Wage Gains and Welfare Gains from Self-Selection Models,” *The Review of Economics and Statistics*, 69, 42–49.
- BLAU, D. AND J. CURRIE (2006): “Pre-School, Day Care, and After School Care: Who’s Minding the Kids?” *Handbook of the Economics of Education*, 2, 1163–1278.
- BLAU, D. M. (1999): “The Effect of Income on Child Development,” *Review of Economics and Statistics*, 81, 261–276.
- (2001): *The Child Care Problem: An Economic Analysis*, Russell Sage Foundation.
- BRADLEY, R. ET AL. (2003): “Child Care and Common Communicable Illnesses in Children Aged 37 to 54 Months,” *Archives of Pediatrics & Adolescent Medicine*, 157, 196.
- BRAVE, S. AND T. WALSTRUM (2014): “Estimating Marginal Treatment Effects Using Parametric and Semiparametric Methods,” *Stata Journal*, 14, 191–217.
- BURKHAUSER, R. V. AND J. CAWLEY (2008): “Beyond BMI: The Value of More Accurate Measures of Fatness and Obesity in Social Science Research,” *Journal of Health Economics*, 27, 519–529.
- CAMPBELL, F., G. CONTI, J. J. HECKMAN, S. H. MOON, R. PINTO, E. PUNGELLO, AND Y. PAN (2014): “Early Childhood Investments Substantially Boost Adult Health,” *Science*, 343, 1478–1485.
- CARD, D. (1999): “The Causal effect of Education on Earnings,” *Handbook of Labor Economics*, 3, 1801–1863.
- (2001): “Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems,” *Econometrica*, 69, 1127–1160.
- CARNEIRO, P. AND R. GINJA (2008): “Preventing Behavior Problems in Childhood and Adolescence: Evidence from Head Start,” *Working Paper, University College London*.

- CARNEIRO, P. AND J. J. HECKMAN (2002): “The Evidence on Credit Constraints in Post-Secondary Schooling\*,” *The Economic Journal*, 112, 705–734.
- CARNEIRO, P., J. J. HECKMAN, AND E. VYTLACIL (2003): “Understanding what Instrumental Variables Estimate: Estimating Marginal and Average Returns to Education,” *University of Chicago, The American Bar Foundation and Stanford University*, July, 19.
- CARNEIRO, P. M., J. J. HECKMAN, AND E. VYTLACIL (2011): “Estimating Marginal Returns to Education,” *American Economic Review*, 101, 2754–81.
- CASCIO, E. U. (2009): “Do Investments in Universal Early Education Pay Off? Long-term Effects of Introducing Kindergartens into Public Schools,” *NBER Working Paper No. 14951*.
- CAWLEY, J. (2004): “The Impact of Obesity on Wages,” *Journal of Human Resources*, 39, 451–474.
- CAWLEY, J. AND F. LIU (2012): “Maternal Employment and Childhood Obesity: A Search for Mechanisms in Time Use Data,” *Economics & Human Biology*, 10, 352–364.
- CAWLEY, J. AND C. K. SPIESS (2008): “Obesity and Skill Attainment in Early Childhood,” *Economics & Human Biology*, 6, 388–397.
- CHETTY, R., J. FRIEDMAN, N. HILGER, E. SAEZ, D. SCHANZENBACH, AND D. YAGAN (2011): “How does your Kindergarten Classroom Affect your Earnings? Evidence from Project Star,” *The quarterly journal of economics*, 126, 1593–1660.
- CLASSEN, T. AND C. HOKAYEM (2005): “Childhood Influences on Youth Obesity,” *Economics & Human Biology*, 3, 165–187.
- CONTI, G., J. HECKMAN, AND S. URZUA (2010): “The Education-Health Gradient,” *The American Economic Review*, 100, 234–238.
- CUNHA, F. AND J. J. HECKMAN (2007): “The Technology of Skill Formation,” *The American Economic Review*, 97, 31–47.
- CURRIE, J. (2009): “Healthy, Wealthy, and Wise: Socioeconomic Status, Poor Health in Childhood, and Human Capital Development,” *Journal of Economic Literature*, 47, 87–122.
- CURRIE, J. AND D. ALMOND (2011): “Human Capital Development Before Age Five,” *Handbook of Labor Economics*, 4, 1315–1486.
- CURRIE, J. AND W. LIN (2007): “Chipping Away at Health: More on the Relationship Between Income and Child Health,” *Health Affairs*, 26, 331–344.

- CURRIE, J., M. STABILE, P. MANIVONG, AND L. L. ROOS (2010): “Child Health and Young Adult Outcomes,” *Journal of Human Resources*, 45, 517–548.
- CURRIE, J. AND D. THOMAS (1995): “Does Head Start Make a Difference?” *American Economic Review*, 85, 341–364.
- DANIELZIK, S., M. CZERWINSKI-MAST, K. LANGNÄSE, B. DILBA, AND M. MÜLLER (2004): “Parental Overweight, Socioeconomic Status and High Birth Weight are the Major Determinants of Overweight and Obesity in 5–7 y-old Children: Baseline Data of the Kiel Obesity Prevention Study (KOPS),” *International Journal of Obesity*, 28, 1494–1502.
- DATTA GUPTA, N. AND M. SIMONSEN (2010): “Non-Cognitive Child Outcomes and Universal High Quality Child Care,” *Journal of Public Economics*, 94, 30–43.
- (2012): “The Effects of Type of Non-Parental Child Care on Pre-Teen Skills and Risky Behavior,” *Economics Letters*, 116, 622–625.
- DAVIS, K. AND K. K. CHRISTOFFEL (1994): “Obesity in Preschool and School-Age Children: Treatment Early and Often May Be Best,” *Archives of Pediatrics & Adolescent Medicine*, 148, 1257–1261.
- DE ONIS, M., M. BLÖSSNER, AND E. BORGI (2010): “Global Prevalence and Trends of Overweight and Obesity among Preschool Children,” *The American Journal of Clinical Nutrition*, 92, 1257–1264.
- DECICCA, P. AND J. SMITH (2013): “The Long-Run Impacts of Early Childhood Education: Evidence from a Failed Policy Experiment,” *Economics of Education Review*, 36, 41–59.
- DEL CARMEN HUERTA, M., W. ADEMA, J. BAXTER, M. CORAK, M. DEDING, M. GRAY, W.-J. HAN, AND J. WALDFOGEL (2011): *Early Maternal Employment and Child Development in Five OECD Countries*, Directorate for Employment, Labour and Social Affairs, OECD.
- DESTATIS (2007): *Kinderbetreuung regional 2006*, Statistisches Bundesamt, Wiesbaden.
- DEUTSCHE RENTENVERSICHERUNG (2012): “3 Rentenzugang,” *Statistik der Deutschen Rentenversicherung, Rentenversicherung in Zeitreihen 2012*.
- D’HONDT, E., B. DEFORCHE, I. DE BOURDEAUDHUIJ, AND M. LENOIR (2009): “Relationship between Motor Skill and Body Mass Index in 5-to 10-Year-Old Children,” *Adapted Physical Activity Quarterly*, 26, 21–37.
- DJI (2008): *Zahlenspiegel 2007 Kindertagesbetreuung im Spiegel der Statistik*, Forschungsverbund Deutschen Jugendinstitut, München.

- D'ONISE, K., J. W. LYNCH, M. G. SAWYER, AND R. A. MCDERMOTT (2010): "Can Preschool Improve Child Health Outcomes? A Systematic Review," *Social Science & Medicine*, 70, 1423–1440.
- DRANGE, N., T. HAVNES, AND A. M. SANDSØR (2012): "Kindergarten for All: Long Run Effects of a Universal Intervention," *Discussion Paper Series, Forschungsinstitut zur Zukunft der Arbeit*.
- DUMAS, C. AND A. LEFRANC (2012): "Early Schooling and Later Outcomes: Evidence from Pre-School Extension in France," in *From Parents to Children - The Intergenerational Transmission of Advantage*, ed. by J. Ermisch, M. Jäntti, and T. M. Smeeding, Russell Sage Foundation.
- DUSTMANN, C., A. RAUTE, AND U. SCHÖNBERG (2012): "Does Universal Child Care Matter? Evidence From a Large Expansion in Pre-School Education," *mimeo, UCL*.
- EBBELING, C. B., D. B. PAWLAK, AND D. S. LUDWIG (2002): "Childhood Obesity: Public-Health Crisis, Common Sense Cure," *The Lancet*, 360, 473–482.
- FELFE, C. AND R. LALIVE (2012): "Early Child Care and Child Development: For Whom it Works and Why," *IZA Discussion Paper No. 7100*.
- (2013): "Does Early Child Care Help or Hinder Children's Development?" *Mimeo*.
- FELFE, C., N. NOLLENBERGER, AND N. RODRIGUEZ PLANAS (2013): "Can't Buy Mommy's Love? Universal Childcare and Children's Long-term Cognitive Development," *CESifo Working Paper: Economics of Education 4069*.
- FINUCANE, M. M., G. A. STEVENS, M. J. COWAN, G. DANAEI, J. K. LIN, C. J. PACIOREK, G. M. SINGH, H. R. GUTIERREZ, Y. LU, A. N. BAHALIM, ET AL. (2011): "National, Regional, and Global Trends in Body-Mass Index since 1980: Systematic Analysis of Health Examination Surveys and Epidemiological Studies with 960 Country-Years and 9· 1 million participants," *The Lancet*, 377, 557–567.
- FITZPATRICK, M. D. (2008): "Starting School at Four: The Effect of Universal Pre-Kindergarten on Children's Academic Achievement," *The BE Journal of Economic Analysis & Policy*, 8, 1–40.
- FITZSIMONS, E. AND M. VERA-HERNANDEZ (2014): "Food for Thought? Breastfeeding and Child Development," *Working Paper No. 14-04*.
- FREEDMAN, D. S., L. K. KHAN, M. K. SERDULA, W. H. DIETZ, S. R. SRINIVASAN, AND G. S. BERENSON (2005): "The Relation of Childhood BMI to Adult Adiposity: The Bogalusa Heart Study," *Pediatrics*, 115, 22–27.
- FRISVOLD, D. E. (2006): "Head Start Participation and Childhood Obesity," *Vanderbilt University Economics Working Paper No. 06-WG01*.

- FRISVOLD, D. E. AND J. C. LUMENG (2011): “Expanding Exposure Can Increasing the Daily Duration of Head Start Reduce Childhood Obesity?” *Journal of Human Resources*, 46, 373–402.
- FRY, J. AND W. FINLEY (2005): “The Prevalence and Costs of Obesity in the EU,” *Proceedings of the Nutrition Society*, 64, 359–362.
- GARCES, E., J. CURRIE, AND D. THOMAS (2002): “Longer-Term Effects of Head Start,” *American economic review*, 92, 999–1012.
- GOODMAN, E., B. R. HINDEN, AND S. KHANDELWAL (2000): “Accuracy of Teen and Parental Reports of Obesity and Body Mass Index,” *Pediatrics*, 106, 52–58.
- GOODMAN, R. (1997): “The Strengths and Difficulties Questionnaire: A Research Note,” *Journal of Child Psychology and Psychiatry*, 38, 581–586.
- GORMLEY JR, W. T., T. GAYER, D. PHILLIPS, AND B. DAWSON (2005): “The Effects of Universal Pre-K on Cognitive Development.” *Developmental Psychology*, 41, 872.
- GROSSMAN, M. (1972): “On the Concept of Health Capital and the Demand for Health,” *The Journal of Political Economy*, 80, 223–255.
- HANK, K., M. KREYENFELD, AND C. SPIESS (2004): “Kinderbetreuung und Fertilität in Deutschland,” *Zeitschrift für Soziologie*, 33, 228–244.
- HAVNES, T. AND M. MOGSTAD (2011): “No Child Left Behind: Subsidized Child Care and Children’s Long-Run Outcomes,” *American Economic Journal: Economic Policy*, 3, 97–129.
- HECKMAN, J. J. (2000): “Policies to Foster Human Capital,” *Research in Economics*, 54, 3–56.
- HECKMAN, J. J. AND D. V. MASTEROV (2007): “The Productivity Argument for Investing in Young Children,” *Applied Economic Perspectives and Policy*, 29, 446–493.
- HECKMAN, J. J., S. H. MOON, R. PINTO, P. A. SAVELYEV, AND A. YAVITZ (2010): “The Rate of Return to the HighScope Perry Preschool Program,” *Journal of Public Economics*, 94, 114–128.
- HECKMAN, J. J., S. URZUA, AND E. VYTLACIL (2006a): “Understanding Instrumental Variables in Models with Essential Heterogeneity,” *The Review of Economics and Statistics*, 88, 389–432.
- (2006b): *Web Supplement to Understanding Instrumental Variables in Models with Essential Heterogeneity: Estimation of Treatment Effects under Essential Heterogeneity.*, [http://jenni.uchicago.edu/underiv/documentation\\_2006\\_03\\_20.pdf](http://jenni.uchicago.edu/underiv/documentation_2006_03_20.pdf) [Online; accessed 30-July-2014].

- HECKMAN, J. J. AND E. VYTLACIL (2001): "Policy-Relevant Treatment Effects," *The American Economic Review*, 91, 107–111.
- HERBST, C. M. AND E. TEKIN (2010): "Child Care Subsidies and Child Development," *Economics of Education Review*, 29, 618–638.
- (2011): "Child Care Subsidies and Childhood Obesity," *Review of Economics of the Household*, 9, 349–378.
- (2012): "The Geographic Accessibility of Child Care Subsidies and Evidence on the Impact of Subsidy Receipt on Childhood Obesity," *Journal of Urban Economics*, 71, 37–52.
- HILLS, A., N. BYRNE, AND J. PARIZKOVA (1998): "Methodological Considerations in the Assessment of Physical Activity and Nutritional Status of Children and Youth," *Medicine and Sport Science*, 43, 155–160.
- IMBENS, G. W. AND J. D. ANGRIST (1994): "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 62, 467–475.
- KLESGES, R. C., L. M. KLESGES, L. H. ECK, AND M. L. SHELTON (1995): "A Longitudinal Analysis of Accelerated Weight Gain in Preschool Children," *Pediatrics*, 95, 126–130.
- KNUDSEN, E. I., J. J. HECKMAN, J. L. CAMERON, AND J. P. SHONKOFF (2006): "Economic, Neurobiological, and Behavioral Perspectives on Building America's Future Workforce," *Proceedings of the National Academy of Sciences*, 103, 10155–10162.
- KONNOPKA, A., M. BÖDEMANN, AND H.-H. KÖNIG (2011): "Health Burden and Costs of Obesity and Overweight in Germany," *The European Journal of Health Economics*, 12, 345–352.
- KREYENFELD, M. AND K. HANK (2000): "Does the Availability of Child Care Influence the Employment of Mothers? Findings from Western Germany," *Population Research and Policy Review*, 19, 317–337.
- KREYENFELD, M., C. K. SPIESS, AND G. G. WAGNER (2002): "Kinderbetreuungspolitik in Deutschland," *Zeitschrift für Erziehungswissenschaft*, 5, 201–221.
- KRÖLL, A. AND R. BORCK (2013): "The Influence of Child Care on Maternal Health and Mother-Child Interaction," *CESifo Working Paper No. 4289*.
- KROMEYER-HAUSCHILD, K. (2005): "Definition, Anthropometric und deutsche Referenzwerte für BMI," in *Adipositas bei Kindern und Jugendlichen*, Springer, 3–15.

- KURTH, B. M., R. DÖLLE, AND H. STOLZENBERG (2013): *KiGGS-Basiserhebung - "Studie zur Gesundheit von Kindern und Jugendlichen in Deutschland"*. Version: 4., Robert Koch Institut, Abteilung für Epidemiologie und Gesundheitsmonitoring, RKI – Robert Koch Institut. Dataset.
- LAMONTAGNE, J. F., P. L. ENGLE, AND M. F. ZEITLIN (1998): "Maternal Employment, Child Care, and Nutritional Status of 12–18-Month-Old Children in Managua, Nicaragua," *Social Science & Medicine*, 46, 403–414.
- LOEB, S., M. BRIDGES, D. BASSOK, B. FULLER, AND R. W. RUMBERGER (2007): "How Much is Too Much? The Influence of Preschool Centers on Children's Social and Cognitive Development," *Economics of Education Review*, 26, 52–66.
- LOKEN, K. V., M. MOGSTAD, AND M. WISWALL (2012): "What Linear Estimators Miss: The Effects of Family Income on Child Outcomes," *American Economic Journal: Applied Economics*, 4, 1–35.
- LUMENG, J., K. GANNON, D. APPUGLIESE, H. CABRAL, AND B. ZUCKERMAN (2004): "Preschool Child Care and Risk of Overweight in 6-to 12-Year-Old Children," *International Journal of Obesity*, 29, 60–66.
- MAHER, E. J., G. LI, L. CARTER, AND D. B. JOHNSON (2008): "Preschool Child Care Participation and Obesity at the Start of Kindergarten," *Pediatrics*, 122, 322–330.
- MAHLER, P. (2008): "Three Essays in Applied Microeconometrics," Ph.D. thesis, University of Zurich.
- MINDLIN, M., R. JENKINS, AND C. LAW (2009): "Maternal Employment and Indicators of Child Health: A Systematic Review in Pre-school Children in OECD Countries," *Journal of Epidemiology and Community Health*, 63, 340–350.
- MOGSTAD, M. AND M. WISWALL (2012): "Testing the Quantity-Quality Model of Fertility: Linearity, Marginal Effects, and Total Effects," *NYU Working Paper*.
- NEWAY, W. K. (1990): "Semiparametric Efficiency Bounds," *Journal of Applied Econometrics*, 5, 99–135.
- NG, M., T. FLEMING, M. ROBINSON, B. THOMSON, N. GRAETZ, C. MARGONO, E. C. MULLANY, S. BIRYUKOV, C. ABBAFATI, S. F. ABERA, ET AL. (2014): "Global, Regional, and National Prevalence of Overweight and Obesity in Children and Adults during 1980–2013: A Systematic Analysis for the Global Burden of Disease Study 2013," *The Lancet*.
- NOBOA-HIDALGO, G. E. AND S. S. URZUA (2012): "The effects of participation in public child care centers: Evidence from Chile," *Journal of Human Capital*, 6, 1–34.
- OECD (2001): *Employment Outlook 2001*, OECD, Paris.

- (2014): *Germany Keeping the Edge: Competitiveness for Inclusive Growth*, OECD“Better Policies”Series, Paris.
- PETER, F. H. (2013): “Day Care (Quality) and Differences in Children’s Health and Personalty Characteristics,” *Conference Paper, 28th Annual Congress of the European Economic Association*.
- PHIPPS, S. A., L. LETHBRIDGE, AND P. BURTON (2006): “Long-run Consequences of Parental Paid Work Hours for Child Overweight Status in Canada,” *Social Science & Medicine*, 62, 977–986.
- RAPP, K., K.-H. SCHICK, H. BODE, AND S. K. WEILAND (2005): “Type of Kindergarten and Other Potential Determinants of Overweight in Pre-School Children,” *Public Health Nutrition*, 8, 642–649.
- RENZ, R. AND B. EGGEN (2004): “Frauen in Europa: Job? Kinder? Oder beides?” *Statistisches Monatsheft Baden-Württemberg*, 11–15.
- ROSSIN, M. (2011): “The Effects of Maternity Leave on Children’s Birth and Infant Health Outcomes in the United States,” *Journal of Health Economics*, 30, 221–239.
- ROY, A. D. (1951): “Some Thoughts on the Distribution of Earnings,” *Oxford Economic Papers*, 3, 135–146.
- RUHM, C. J. (2000): “Parental Leave and Child Health,” *Journal of Health Economics*, 19, 931–960.
- (2004): “Parental Employment and Child Cognitive Development,” *Journal of Human Resources*, 39, 155–192.
- (2008): “Maternal Employment and Adolescent Development,” *Labour Economics*, 15, 958–983.
- SERDULA, M. K., D. IVERY, R. J. COATES, D. S. FREEDMAN, D. F. WILLIAMSON, AND T. BYERS (1993): “Do Obese Children Become Obese Adults? A Review of the Literature,” *Preventive Medicine*, 22, 167–177.
- SHEA, J. (1997): “Instrument Relevance in Multivariate Linear Models: A Simple Measure,” *Review of Economics and Statistics*, 79, 348–352.
- SLAUGHTER, M., T. LOHMAN, R. BOILEAU, C. HORSWILL, R. STILLMAN, M. VAN LOAN, AND D. BEM BEN (1988): “Skinfold Equations for Estimation of Body Fatness in Children and Youth.” *Human biology*, 60, 709.
- SPIESS, C. K., F. BÜCHEL, AND J. R. FRICK (2002): “Kinderbetreuung in West-und Ostdeutschland: Sozioökonomischer Hintergrund entscheidend,” *DIW Wochenbericht*, 69, 518–524.
- SPIESS, C. K., F. BÜCHEL, AND G. G. WAGNER (2003): “Children’s School Placement in Germany: Does Kindergarten Attendance Matter?” *Early Childhood Research Quarterly*, 18, 255–270.

- SPIESS, C. K. AND W. TIETZE (2002): "Qualitätssicherung in Kindertageseinrichtungen," *Zeitschrift für Erziehungswissenschaft*, 5, 139–162.
- STAIGER, D. AND J. H. STOCK (1997): "Instrumental Variables Regression with Weak Instruments," *Econometrica*, 65, 557–586.
- STODDEN, D., S. LANGENDORFER, AND M. A. ROBERTON (2009): "The Association Between Motor Skill Competence and Physical Fitness in Young Adults," *Research Quarterly for Exercise and Sport*, 80, 223–229.
- TAKAHASHI, E., K. YOSHIDA, H. SUGIMORI, M. MIYAKAWA, T. IZUNO, T. YAMAGAMI, AND S. KAGAMIMORI (1999): "Influence Factors on the Development of Obesity in 3-Year-Old Children Based on the Toyama Study," *Preventive Medicine*, 28, 293–296.
- TANAKA, S. (2005): "Parental Leave and Child Health Across OECD Countries," *The Economic Journal*, 115, F7–F28.
- TSAI, A. G., D. F. WILLIAMSON, AND H. A. GLICK (2011): "Direct Medical Cost of Overweight and Obesity in the USA: A Quantitative Systematic Review," *Obesity Reviews*, 12, 50–61.
- VON HINKE KESSLER SCHOLDER, S. (2008): "Maternal Employment and Overweight Children: Does Timing Matter?" *Health Economics*, 17, 889–906.
- VOTRUBA-DRZAL, E., R. L. COLEY, AND P. L. CHASE-LANSDALE (2004): "Child Care and Low-Income Children's Development: Direct and Moderated Effects," *Child development*, 75, 296–312.
- WALDFOGEL, J., W.-J. HAN, AND J. BROOKS-GUNN (2002): "The Effects of Early Maternal Employment on Child Cognitive Development," *Demography*, 39, 369–392.
- WOLL, A., B.-M. KURTH, E. OPPER, A. WORTH, AND K. BÖS (2011): "The 'Motorik-Modul' (MoMo): Physical fitness and physical activity in German children and adolescents," *European Journal of Pediatrics*, 170, 1129–1142.
- WOOLDRIDGE, J. M. (2002): *Econometric Analysis of Cross Section and Panel Data*, The MIT press.
- WROHLICH, K. (2008): "The Excess Demand for Subsidized Child Care in Germany," *Applied Economics*, 40, 1217–1228.

## A Controls

Table A1: Control variables

<b>Child characteristics</b>	
Age	Dummies for each 1/2 of a year
Boy	Dummy, 1= Boy
Height	Size of child in cm
Age	in months, linear, quadratic, cubic
Born before 2000	Dummy, 1=Born before 2000
Malformation	Dummy, 1=Child had malformation after birth
Use Childminder	Dummy, 1=Cared by a childminder
School age	Dummy, 1= Age of child $\geq$ 6
School now***	Dummy, 1=Child is in school now
<b>Parents and Family</b>	
Migrant	Dummy, 1=migrant*
First Child	Dummy, 1=First born
Sibling	Number
Separated Parents*	Dummy, 1=Single/step parent
Age Mother	in years, linear and squared
Age Father	in years, linear and squared
Mother's Training:	Dummies, University or Technical College, Vocational, only school info, Reference: no training/school****
Father's Training:	Dummies, University or Technical College, Vocational, only school info, no info, Reference: no training/school****
Parents Allergy	dummy, 1=parents have allergies
Mother Overweight	Dummy, 1= $>$ 25BMI < 30
Mother Obese	Dummy, 1= $\geq$ 30BMI < 35
Mother Highly Obese	Dummy, 1=BMI $\geq$ 35
Father Overweight	Dummy, 1= $>$ 25BMI < 30
Father Obese	Dummy, 1= $\geq$ 30BMI < 35
Father Highly Obese	Dummy, 1=BMI $\geq$ 35
Breastfed	Dummy, 1 = Child was breastfed
Smoke in Flat	Parents smoke in flat, Dummies: Sometimes, Often
Month	Dummy, 1= Months reported
<b>Regional controls</b>	
Unemployment rate** (county)	linear
Fraction of Working Population** (county)	linear
Fraction of Working Women aged 30-50** (county)	linear
Fraction of Foreigners ** (county)	linear
Income per Capita (pc)** (in 1000) (county)	linear and squared
Population Density** (county)	1000 Inhabitants divided by area size (km <sup>2</sup> )
Fraction of Population age 0-3 ** (county)	linear
City**	Dummies, 4 categories: Small city=5000-20000, Mid. city=20000-100000,
State Dummies	16 dummies
<b>Instrument</b>	
Slot-Child Ration	Available slots for children aged 0-3 divided by number of children age 0-3

\* Child migrated to Germany and at least one parent was not born in Germany or both parents are migrated or have not the German citizenship.

\*\* These variables are also interacted with an East Germany dummy.

\*\*\* Only asked in the questionnaire for children that are seven years old or older.

Source: KiGGS Version 4 & Federal Statistical Office, own calculations.

## B OLS

Table A2: Variation of OLS estimates with current monthly household income

	Body-Mass-Index			Body Fat		Motor Jump
	BMI	Overweight	Obese	BF	BF>p90	
<i>Monthly household income ≥ €3000</i>						
<i>Early Care</i>	0.321*	0.062*	0.015	0.252	0.032	-0.062
	(0.173)	(0.032)	(0.015)	(0.503)	(0.022)	(0.119)
Observations	754	754	754	747	747	745
<i>Monthly household income &lt; €3000</i>						
<i>Early Care</i>	-0.286*	-0.059**	-0.021	-0.800*	-0.029	0.090
	(0.171)	(0.027)	(0.017)	(0.470)	(0.027)	(0.070)
Observations	1618	1618	1618	1601	1601	1590
<i>Monthly household income €1750-3000</i>						
<i>Early Care</i>	-0.312	-0.060**	-0.034	-0.612	-0.018	0.197**
	(0.190)	(0.030)	(0.021)	(0.525)	(0.033)	(0.093)
Observations	1117	1117	1117	1099	1099	1091
<i>Monthly household income &lt; €1750</i>						
<i>Early Care</i>	-0.920**	-0.149**	-0.057	-3.069***	-0.154***	0.111
	(0.440)	(0.065)	(0.038)	(1.030)	(0.051)	(0.117)
Observations	348	348	348	349	349	346

*Notes:* All estimates are based on IV estimations controlling for the set of variables defined in ‘Child characteristic’, ‘Parents and Family’ and ‘Regional controls’ as described in table A1.

Inference is based on standard errors that allow for clustering at the county level.

\* / \*\* / \*\*\*: Significance at the 10-/5-/1-percent level.

*Source:* KiGGS Version 4 & Federal Statistical Office, own calculations.

Table A3: Variation of OLS estimates with paternal BMI

	<b>Body-Mass-Index</b>			<b>Body Fat</b>		<b>Motor</b>
	BMI	Overweight	Obese	BF	BF>p90	Jump
<i>BMI father ≤ 25</i>						
<i>Early Care</i>	-0.089 (0.177)	-0.018 (0.029)	-0.011 (0.021)	-0.489 (0.442)	-0.012 (0.029)	0.112 (0.078)
Observations	1354	1354	1354	1342	1342	1333
<i>BMI father &gt; 25</i>						
<i>Early Care</i>	-0.024 (0.163)	0.004 (0.025)	0.005 (0.012)	-0.168 (0.457)	0.010 (0.024)	-0.020 (0.079)
Observations	1178	1178	1178	1167	1167	1163

*Notes:* All estimates are based on IV estimations controlling for the set of variables defined in ‘Child characteristic’, ‘Parents and Family’ and ‘Regional controls’ as described in table A1.

Inference is based on standard errors that allow for clustering at the county level.

\* / \*\* / \*\*\*: Significance at the 10-/5-/1-percent level.

*Source:* KiGGS Version 4 & Federal Statistical Office, own calculations.

## C Propensity Score

Table A4: First stage of TSLS (OLS)

	<i>Early Care</i>	
Slot-Child Ratio	1.071***	(0.314)
Boy	-0.018	(0.016)
Age(month)	-1.001	(0.648)
Age(month) <sup>2</sup>	0.136	(0.089)
Age(month) <sup>3</sup>	-0.006	(0.004)
Born before 2000	-0.017	(0.031)
Low Birth Weight	-0.016	(0.029)
Malformation	-0.063***	(0.020)
Height (cm)	0.003	(0.002)
First Child	0.027	(0.028)
No Siblings	-0.038***	(0.011)
Separated Parent	0.061**	(0.025)
Migrant	-0.039	(0.029)
Age Mother	0.022	(0.019)
Age Mother <sup>2</sup>	-0.000	(0.000)
Age Father	-0.020	(0.017)
Age Father <sup>2</sup>	0.000	(0.000)
Mother Uni/College	0.094**	(0.041)
Mother Vocational	0.028	(0.030)
Mother in Training	0.169**	(0.084)
Mother only school info	0.148	(0.115)
Father Uni/College	-0.061	(0.048)
Father Vocational	-0.065	(0.043)
Father in Training	-0.080	(0.117)
Father only school info	-0.160	(0.112)
Father no info	-0.005	(0.117)
Parent Allergy	0.008	(0.014)
Mother Overweight	-0.005	(0.015)
Mother Obese	0.006	(0.025)
Mother Highly Obese	-0.044	(0.037)
Father Overweight	0.004	(0.015)
Father Obese	0.019	(0.022)
Father Highly Obese	0.044	(0.061)
Smoke sometimes	0.016	(0.021)
Smoke regularly	0.048**	(0.021)
Breastfed	0.049***	(0.017)
Breastfed no info	-0.154	(0.113)
Month reported	0.098***	(0.018)
Use Childminder	0.450***	(0.045)
School-age	0.033	(0.040)
School now	-0.043	(0.032)
Unempl. Rate	-0.008	(0.007)
Employed(frac.)	-0.002	(0.002)
Empl. women age 30-50(frac.)	0.001	(0.001)
Foreigners(frac.)	-0.690	(0.506)
Income(pc) in ths	-0.086	(0.062)
Income(pc) <sup>2</sup>	0.003*	(0.002)
Population Density (ths/km <sup>2</sup> )	0.068***	(0.024)
Small City	0.006	(0.025)
Mid. City	0.004	(0.027)
Big City	0.026	(0.044)
Pop.0-3(frac.)	-0.004	(0.047)
Unempl. Rate * East	0.018*	(0.010)
Employed (frac.)* East	-0.003	(0.004)
Empl. women age 30-50 (frac.)* East	0.001	(0.002)
Foreigners (frac.)* East	1.379	(4.306)
Income(pc)* East	0.137	(0.392)
Income(pc)sq. * East	-0.004	(0.015)

Table A4: First stage of TSLS (OLS) (continued)

Density * East	-0.036	(0.075)
Small City * East	-0.084	(0.055)
Mid. City * East	-0.046	(0.057)
Big City * East	-0.156	(0.132)
Pop.0-3(frac.)* East	0.055	(0.119)
STATE FIXED EFFECTS		YES
Constant	2.925*	(1.707)
Observations	2377	

*Notes:* Estimates are based on an OLS estimations controlling for the set of variables defined in ‘Child characteristic’, ‘Parents and Family’ and ‘Regional controls’ as described in table A1.

Inference is based on standard errors that allow for clustering at the school district level.

\* / \*\* / \*\*\*. Significance at the 10-/5-/1-percent level.

*Source:* KiGGS Version 4 & Federal Statistical Office, own calculations.

## D Full results

Table A5: NTSLS - Full results

	Body-Mass-Index			Body Fat		Motor
	BMI	Overweight	Obese	BF	BF>p90	Jump
<i>Early Care</i>	-1.064**	-0.130**	-0.094**	-2.404**	-0.116*	0.467**
	(0.443)	(0.065)	(0.042)	(1.105)	(0.061)	(0.219)
Boy	-0.133	-0.009	-0.002	-0.305	-0.020	-0.042
	(0.084)	(0.013)	(0.009)	(0.219)	(0.012)	(0.040)
Age(month)	-3.447	-0.284	-0.856**	-9.517	-0.126	0.513
	(4.066)	(0.615)	(0.379)	(11.037)	(0.581)	(1.954)
Age(month) <sup>2</sup>	0.382	0.028	0.115**	1.035	0.002	-0.121
	(0.574)	(0.086)	(0.053)	(1.569)	(0.080)	(0.270)
Age(month) <sup>3</sup>	-0.018	-0.001	-0.005**	-0.046	0.000	0.007
	(0.026)	(0.004)	(0.002)	(0.073)	(0.004)	(0.012)
Born before 2000	-0.126	-0.010	-0.006	0.110	0.009	0.088
	(0.153)	(0.026)	(0.016)	(0.394)	(0.026)	(0.089)
Low Birth Weight	0.193	0.021	-0.002	0.426	0.009	0.087
	(0.164)	(0.027)	(0.017)	(0.419)	(0.025)	(0.077)
Malformation	0.035	-0.008	-0.006	0.376	0.006	-0.068
	(0.134)	(0.021)	(0.013)	(0.367)	(0.020)	(0.067)
Height (cm)	0.124***	0.016***	0.007***	0.292***	0.013***	-0.008**
	(0.009)	(0.001)	(0.001)	(0.023)	(0.001)	(0.004)
First Child	0.016	0.011	-0.009	0.194	-0.019	0.054
	(0.125)	(0.021)	(0.013)	(0.341)	(0.019)	(0.064)
No Siblings	-0.074	-0.020***	-0.004	-0.273**	-0.023***	0.003
	(0.050)	(0.007)	(0.005)	(0.128)	(0.006)	(0.026)
Separated Parent	0.128	0.010	0.007	0.318	0.030	0.026
	(0.157)	(0.026)	(0.015)	(0.422)	(0.023)	(0.071)
Migrant	0.252	0.050*	0.027	1.193**	0.059**	0.056
	(0.186)	(0.028)	(0.021)	(0.473)	(0.028)	(0.086)
Age Mother	-0.000	-0.018	0.006	-0.145	0.000	0.101**
	(0.097)	(0.018)	(0.011)	(0.249)	(0.016)	(0.044)
Age Mother <sup>2</sup>	-0.000	0.000	-0.000	0.002	0.000	-0.001**
	(0.001)	(0.000)	(0.000)	(0.003)	(0.000)	(0.001)
Age Father	-0.063	0.008	-0.002	-0.053	-0.002	-0.020
	(0.086)	(0.011)	(0.010)	(0.223)	(0.011)	(0.033)
Age Father <sup>2</sup>	0.001	-0.000	0.000	0.001	0.000	0.000
	(0.001)	(0.000)	(0.000)	(0.003)	(0.000)	(0.000)
Mother Uni/College	0.315	-0.001	0.024	0.440	-0.037	0.027
	(0.233)	(0.035)	(0.026)	(0.596)	(0.034)	(0.106)
Mother Vocational	0.153	0.007	0.011	0.440	-0.014	0.062
	(0.199)	(0.029)	(0.022)	(0.513)	(0.030)	(0.085)
Mother in Training	-0.268	0.033	0.089	0.238	-0.002	-0.134
	(0.446)	(0.077)	(0.066)	(1.165)	(0.067)	(0.180)
Mother only school info	0.188	-0.059	0.027	-0.139	0.041	-0.062
	(0.547)	(0.068)	(0.058)	(1.233)	(0.072)	(0.196)
Father Uni/College	-0.385	-0.020	-0.025	-0.973	-0.063	0.184
	(0.281)	(0.040)	(0.031)	(0.796)	(0.040)	(0.120)
Father Vocational	-0.302	-0.026	-0.014	-0.695	-0.047	0.140
	(0.268)	(0.038)	(0.030)	(0.769)	(0.039)	(0.110)
Father in Training	-0.618	-0.010	-0.050	-2.063*	-0.127***	0.028
	(0.477)	(0.078)	(0.038)	(1.171)	(0.048)	(0.264)
Father only school info	-0.069	0.068	-0.014	-0.119	-0.075	-0.189
	(0.650)	(0.099)	(0.060)	(1.570)	(0.082)	(0.215)
Father no info	0.086	0.064	0.036	-0.218	0.009	-0.303
	(0.768)	(0.129)	(0.097)	(1.805)	(0.113)	(0.215)
Parent Allergy	-0.129	-0.001	-0.007	-0.291	0.005	0.118***
	(0.083)	(0.013)	(0.008)	(0.221)	(0.012)	(0.041)
Mother Overweight	0.445***	0.033**	0.026**	0.982***	0.033**	-0.080
	(0.104)	(0.016)	(0.011)	(0.282)	(0.015)	(0.049)
Mother Obese	1.260***	0.170***	0.055**	3.062***	0.133***	-0.038

Table A5: NTSLS - Full results (continued)

	(0.194)	(0.034)	(0.023)	(0.509)	(0.031)	(0.076)
Mother Highly Obese	1.235***	0.182***	0.073*	2.951***	0.104**	-0.401***
	(0.355)	(0.054)	(0.044)	(0.815)	(0.051)	(0.112)
Father Overweight	0.378***	0.055***	0.020**	0.845***	0.033***	-0.041
	(0.086)	(0.013)	(0.008)	(0.231)	(0.012)	(0.044)
Father Obese	0.836***	0.112***	0.070***	1.561***	0.076***	-0.145**
	(0.179)	(0.028)	(0.021)	(0.447)	(0.026)	(0.073)
Father Highly Obese	1.352***	0.198***	0.081*	2.753**	0.118*	0.034
	(0.433)	(0.067)	(0.048)	(1.234)	(0.063)	(0.172)
Smoke sometimes	0.504***	0.066***	0.027*	1.246***	0.044**	-0.077
	(0.150)	(0.022)	(0.016)	(0.382)	(0.021)	(0.066)
Smoke regularly	0.602***	0.088***	0.064***	1.212***	0.040*	-0.078
	(0.160)	(0.025)	(0.020)	(0.414)	(0.023)	(0.065)
Breastfed	-0.086	-0.006	-0.018	-0.526	-0.019	-0.042
	(0.122)	(0.019)	(0.014)	(0.326)	(0.018)	(0.057)
Breastfed no info	-0.859	-0.111	-0.122***	-2.338	-0.105	0.132
	(0.677)	(0.131)	(0.032)	(1.826)	(0.123)	(0.173)
Month reported	0.156	0.011	0.007	0.372	0.003	0.016
	(0.103)	(0.017)	(0.011)	(0.266)	(0.015)	(0.051)
Use Childminder	0.440	0.041	0.031	0.757	0.028	-0.130
	(0.270)	(0.039)	(0.024)	(0.629)	(0.035)	(0.133)
School-age	0.122	0.014	0.011	0.379	0.016	0.201*
	(0.189)	(0.031)	(0.018)	(0.469)	(0.032)	(0.108)
School now	0.020	0.019	-0.004	-0.198	-0.008	0.087
	(0.194)	(0.032)	(0.020)	(0.529)	(0.029)	(0.099)
Unempl. Rate	-0.105**	-0.012*	-0.005	-0.337***	-0.013**	0.016
	(0.041)	(0.006)	(0.004)	(0.107)	(0.006)	(0.020)
Employed(frac.)	0.017	0.002	0.000	0.002	-0.002	-0.004
	(0.016)	(0.003)	(0.002)	(0.042)	(0.002)	(0.008)
Empl. women age 30-50(frac.)	-0.006	-0.001	-0.000	0.012	0.002	0.005
	(0.012)	(0.002)	(0.001)	(0.030)	(0.002)	(0.006)
Foreigners(frac.)	2.358	0.187	0.000	10.708	0.210	2.514*
	(2.685)	(0.460)	(0.260)	(7.279)	(0.408)	(1.377)
Income(pc) in ths	-0.232	0.014	-0.046	-0.043	0.018	-0.181
	(0.449)	(0.073)	(0.047)	(1.233)	(0.066)	(0.256)
Income(pc) <sup>2</sup>	0.004	-0.001	0.001	-0.007	-0.001	0.006
	(0.014)	(0.002)	(0.001)	(0.038)	(0.002)	(0.008)
Population Density (ths/km <sup>2</sup> )	0.045	-0.019	0.001	0.052	-0.032	-0.097
	(0.188)	(0.030)	(0.019)	(0.536)	(0.029)	(0.089)
Small City	0.243	0.048**	0.051***	0.112	0.008	0.136*
	(0.152)	(0.023)	(0.015)	(0.396)	(0.022)	(0.076)
Mid. City	0.171	0.037	0.037**	0.019	0.018	0.074
	(0.167)	(0.026)	(0.016)	(0.442)	(0.025)	(0.090)
Big City	0.273	0.104*	0.071*	0.268	0.114**	-0.146
	(0.384)	(0.060)	(0.040)	(1.028)	(0.055)	(0.199)
Pop.0-3(frac.)	-0.441	-0.056	-0.009	-1.753**	-0.020	0.082
	(0.299)	(0.046)	(0.032)	(0.807)	(0.046)	(0.166)
Unempl. Rate * East	0.141**	0.021**	0.006	0.524***	0.028***	-0.091***
	(0.056)	(0.008)	(0.006)	(0.149)	(0.008)	(0.026)
Employed (frac.)* East	-0.005	0.001	-0.002	0.030	0.004	0.003
	(0.027)	(0.004)	(0.003)	(0.065)	(0.004)	(0.012)
Empl. women age 30-50 (frac.)* East	0.007	-0.000	0.002	-0.006	-0.003	-0.005
	(0.016)	(0.002)	(0.002)	(0.040)	(0.002)	(0.007)
Foreigners (frac.)* East	8.139	2.287	0.613	47.747	3.910*	-8.234
	(16.381)	(2.494)	(1.428)	(46.412)	(2.325)	(9.094)
Income(pc)* East	-2.923	-0.568*	-0.191	-7.030	0.231	-0.098
	(2.026)	(0.332)	(0.208)	(5.556)	(0.282)	(0.925)
Income(pc)sq. * East	0.119	0.022*	0.009	0.291	-0.009	-0.004
	(0.078)	(0.013)	(0.008)	(0.214)	(0.011)	(0.036)
Density * East	0.035	-0.034	-0.032	0.187	-0.021	0.357**
	(0.362)	(0.058)	(0.031)	(0.995)	(0.049)	(0.170)
Small City * East	-0.404	-0.040	-0.077***	0.425	-0.016	0.069

Table A5: NTSLS - Full results (continued)

	(0.264)	(0.040)	(0.026)	(0.695)	(0.038)	(0.120)
Mid. City * East	-0.929***	-0.093**	-0.080***	-1.550**	-0.102**	0.063
	(0.263)	(0.042)	(0.027)	(0.709)	(0.040)	(0.137)
Big City * East	-1.458**	-0.122	-0.120**	-4.020**	-0.223***	0.312
	(0.589)	(0.098)	(0.053)	(1.636)	(0.084)	(0.308)
Pop.0-3(frac.)* East	0.174	0.083	-0.033	0.582	-0.023	-0.170
	(0.580)	(0.089)	(0.059)	(1.552)	(0.083)	(0.275)
STATE DUMMIES	YES	YES	YES	YES	YES	YES
Constant	2.077	-0.458	1.727*	8.423	-0.681	-0.434
	(10.725)	(1.616)	(1.007)	(28.759)	(1.570)	(5.357)
Observations	2372	2372	2372	2348	2348	2335

*Notes:* All estimates are based on IV estimations controlling for the set of variables defined in 'Child characteristic', 'Parents and Family' and 'Regional controls' as described in table A1. Inference is based on standard errors that allow for clustering at the county level.

\* / \*\* / \*\*\*: Significance at the 10-/5-/1-percent level.

*Source:* KiGGS Version 4 & Federal Statistical Office, own calculations.

## E Robustness

Table A6: NTSLS - The role of control variables

	Body-Mass-Index			Body Fat		Motor
	BMI	Overweight	Obese	BF	BF>p90	Jump
<i>Only child characteristic and parental education</i>						
<i>Early Care</i>	-0.794 (1.044)	-0.099 (0.142)	-0.072 (0.088)	0.300 (2.762)	-0.073 (0.106)	0.776 (0.480)
<i>No regional</i>						
<i>Early Care</i>	-0.887** (0.441)	-0.107 (0.066)	-0.097** (0.041)	-1.414 (1.213)	-0.085* (0.051)	0.520** (0.243)
<i>Standard +plus interactions of socioeconomic and health with East</i>						
<i>Early Care</i>	-1.044*** (0.376)	-0.124** (0.060)	-0.094** (0.040)	-2.653** (1.033)	-0.124** (0.054)	0.519** (0.221)
<i>More regional <sup>a</sup></i>						
<i>Early Care</i>	-0.914** (0.378)	-0.103* (0.057)	-0.100*** (0.037)	-2.470** (1.081)	-0.122** (0.053)	0.601*** (0.215)
<i>Standard plus dummies for administrative regions ('Regierungsbezirk')</i>						
<i>Early Care</i>	-0.719* (0.415)	-0.068 (0.065)	-0.072* (0.037)	-1.239 (1.080)	-0.049 (0.056)	0.446* (0.226)
Observations	2372	2372	2372	2348	2348	2335
<i>No city states (Hamburg, Bremen, Berlin)</i>						
<i>Early Care</i>	-1.129** (0.463)	-0.155** (0.074)	-0.092* (0.047)	-2.148* (1.118)	-0.095* (0.054)	0.470* (0.250)
Observations	2080	2080	2080	2056	2056	2046

Notes: All estimates are based on IV estimations controlling for the set of variables defined in 'Child characteristic', 'Parents and Family' and 'Regional controls' as described in table A1.

Inference is based on standard errors that allow for clustering at the county level.

<sup>a</sup>: GDP (linear and squared), Fraction of population age 30 to 50, population size (interacted with East), Income pc in 2002, average household income and BMI of mother, fathers and children two years old or younger in at the sample point .

\* / \*\* / \*\*\*: Significance at the 10-/5-/1-percent level.

Source: KiGGS Version 4 & Federal Statistical Office, own calculations.

Table A7: NTSLs- Modeling stage 0: Flexible specification of the instrument

	<b>Body-Mass-Index</b>			<b>Body Fat</b>		<b>Motor</b>
	BMI	Overweight	Obese	BF	BF>p90	Jump
<i>Slot-Child Ratio 1998</i>						
<i>Early Care</i>	-1.119*** (0.392)	-0.131** (0.061)	-0.099** (0.039)	-2.146* (1.101)	-0.108** (0.053)	0.432* (0.225)
F-stat.(NTSLs): 213.69, F-stat.(TSLs): 29.68 , Part. R-sq.(TSLs): 0.013						
<i>Plus Interaction East</i>						
<i>Early Care</i>	-0.937** (0.405)	-0.109* (0.064)	-0.078* (0.040)	-2.171** (1.082)	-0.105* (0.056)	0.452** (0.220)
F-stat.(NTSLs): 205.87, F-stat.(TSLs): 11.44 , Part. R-sq.(TSLs): 0.010						
<i>Plus Slot-Child Ratio Squared</i>						
<i>Early Care</i>	-0.874** (0.420)	-0.114* (0.063)	-0.080* (0.042)	-2.315** (1.121)	-0.112** (0.056)	0.461** (0.219)
Observations	2372	2372	2372	2348	2348	2335
F-stat.(NTSLs): 205.61, F-stat.(TSLs): 13.36 , Part. R-sq.(TSLs): 0.012						

Notes: All estimates are based on IV estimations controlling for the set of variables defined in 'Child characteristic', 'Parents and Family' and 'Regional controls' as described in table A1. Inference is based on standard errors that allow for clustering at the county level.

\* / \*\* / \*\*\*: Significance at the 10-/5-/1-percent level.

Source: KiGGS Version 4 & Federal Statistical Office, own calculations.

Table A8: NTSLs- The role of the threshold definition

	<b>Body-Mass-Index</b>			<b>Body Fat</b>		<b>Motor</b>
	BMI	Overweight	Obese	BF	BF>p90	Jump
<i>Later Cutoff (&lt; 3 years = 36 months)</i>						
<i>Early Care (35)</i>	-0.891** (0.372)	-0.115* (0.060)	-0.070* (0.040)	-2.085** (0.982)	-0.077 (0.050)	0.419* (0.215)
Observations	2372	2372	2372	2348	2348	2335
F-stat.(NTSLs): 196.14, F-stat.(TSLs): 17.21 , Part. R-sq.(TSLs): 0.007						
<i>Different rounding (25 to 35 month downward rounded to ≤ 2 years)</i>						
<i>Early Care (R2)</i>	-1.407*** (0.496)	-0.181** (0.076)	-0.101** (0.049)	-2.843** (1.366)	-0.150** (0.069)	0.605** (0.281)
Observations	2372	2372	2372	2348	2348	2335
F-stat.(NTSLs): 119.53, F-stat.(TSLs): 16.86 , Part. R-sq.(TSLs): 0.007						

Notes: All estimates are based on IV estimations controlling for the set of variables defined in 'Child characteristic', 'Parents and Family' and 'Regional controls' as described in table A1. Inference is based on standard errors that allow for clustering at the county level.

\* / \*\* / \*\*\*: Significance at the 10-/5-/1-percent level.

Source: KiGGS Version 4 & Federal Statistical Office, own calculations.

Table A9: NTSLs- More robustness

	Body-Mass-Index			Body Fat		Motor
	BMI	Overweight	Obese	BF	BF>p90	Jump
<i>Allow for children who were only cared by a childminder</i>						
<i>Early Care</i>	-0.975** (0.376)	-0.101* (0.057)	-0.080** (0.038)	-2.166** (1.012)	-0.096* (0.052)	0.356* (0.214)
Observations	2487	2487	2487	2460	2460	2450
F-stat.(NTSLs): 232.47 , F-stat.(TSLs): 23.51, Part. R-sq.(TSLs): 0.010						
<i>Allow for children born in 1995</i>						
<i>Early Care</i>	-0.790* (0.443)	-0.082 (0.057)	-0.082** (0.037)	-2.019* (1.092)	-0.100* (0.054)	0.357 (0.232)
Observations	2733	2733	2733	2702	2702	2691
F-stat.(NTSLs): 193.92 , F-stat.(TSLs): 25.07, Part. R-sq.(TSLs): 0.009						
<i>Allow for children who started daycare later than with 54 months</i>						
<i>Early Care</i>	-0.903** (0.421)	-0.111* (0.064)	-0.074* (0.042)	-2.066* (1.133)	-0.102* (0.057)	0.392* (0.219)
Observations	2427	2427	2427	2402	2402	2388
F-stat.(NTSLs): 203.46, F-stat.(TSLs): 18.02, Part. R-sq.(TSLs): 0.008						

*Notes:* All estimates are based on IV estimations controlling for the set of variables defined in ‘Child characteristic’, ‘Parents and Family’ and ‘Regional controls’ as described in table A1. Inference is based on standard errors that allow for clustering at the county level.

\* / \*\* / \*\*\*: Significance at the 10-/5-/1-percent level.

*Source:* KiGGS Version 4 & Federal Statistical Office, own calculations.

Table A10: NTSLs - School age

	Body-Mass-Index			Body Fat		Motor
	BMI	Overweight	Obese	BF	BF>p90	Jump
<i>NTSLs-Age 6-9</i>						
<i>Early Care</i>	-1.254** (0.529)	-0.155** (0.074)	-0.132*** (0.048)	-3.571*** (1.332)	-0.189*** (0.066)	0.435* (0.234)
Observations	1803	1803	1803	1785	1785	1775
F-stat.(NTSLs): 178.16 , F-stat.(TSLs): 23.20 , Part. R-sq.(TSLs): 0.013						

*Notes:* All estimates are based on IV estimations controlling for the set of variables defined in ‘Child characteristic’, ‘Parents and Family’ and ‘Regional controls’ as described in table A1. Inference is based on standard errors that allow for clustering at the county level.

\* / \*\* / \*\*\*: Significance at the 10-/5-/1-percent level.

*Source:* KiGGS Version 4 & Federal Statistical Office, own calculations.

## F Other outcomes

Measures of socio-emotional development are reported by the parents in form of the Strengths and Difficulties Questionnaire (SDQ). The SDQ score is based on a battery of questions answered by the parent. The scale ranges from 0 to 10, with a value of 10 indicating a very problematic child in a specific area. These areas are emotional, conduct, hyperactivity or peer problems. THE SDQ value reported here is the sum of the score in the four mentioned areas and as could have a value between 0 and 40. This is a standard behavioral measure in the child development literature (Goodman 1997).

In the motor test section, forward bends were used for the assessment of flexibility. The lowest point reached by the fingertips while standing on a box with legs extended was recorded. Single leg stand was used for assessing gross motor coordination during static precision tasks. Subjects were asked to stand on their dominant leg for one minute with their eyes open, and the number of floor contacts with the contralateral limb was recorded. Fine motor coordination during precision tasks was assessed using the sorting pens tests. In pen test the children were then asked to move 25 pens from a pen holder into wholes at the edge of a board as quickly as possible. This time was counted.<sup>66</sup> Reaction time was assessed using a computer aided reaction test. Subjects were asked to react as quickly as possible to 14 color changes of a traffic light, and the seven fastest times were averaged.

These tests are also standardized to mean zero and standard variation of one within each age-gender group. The tests are scaled such that a higher value indicates a more desirable score.

Table A11: NTSLs - Other outcomes

	BMI<p10	Bend	One Leg	Pen	React	SDQ	Emotion	Conduct	Hyper	Peer
<i>Early Care</i>	0.040 (0.065)	0.195 (0.218)	-0.224 (0.206)	-0.222 (0.203)	-0.260 (0.185)	-0.794 (1.018)	0.233 (0.355)	-0.587* (0.327)	-0.063 (0.459)	-0.365 (0.277)
Observations	2372	2337	2327	2343	2345	2369	2371	2371	2371	2372

*Notes:* All estimates are based on IV estimations controlling for the set of variables defined in 'Child characteristic', 'Parents and Family' and 'Regional controls' as described in table A1. Inference is based on standard errors that allow for clustering at the school district level. \*/ \*\*/ \*\*\*: Significance at the 10-/5-/1-percent level.

*Source:* KiGGS Version 4 & Federal Statistical Office, own calculations.

<sup>66</sup>The tracing lines test is an additional fine motor test. In this test the children were asked to trace a groove on a board using a pen without touching the rims of the groove. However, the only variable provided in the data is the average time per mistake. This is a ratio, which is an index without very clear interpretation. We prefer not to discuss this outcome.

## G Distributions

Table A12: Center-based care and Childminder

Center-Based Care	Childminder		Total
	No	Yes	
No	264	3,316	3,580
Yes	61	0	61
Total	325	3,316	3,641

Source: KiGGS Version 4, own calculations.

Table A13: Distribution of outcomes - Mean in full sample

	Early Care		Total
	Yes	No	
BMI	-0.02	-0.12	-0.08
Overweight	0.12	0.12	0.12
Obese	0.04	0.05	0.04
BF	-0.05	-0.27	-0.19
BF>p90	0.10	0.10	0.10
Jump	-0.02	0.01	0.00

Source: KiGGS Version 4, own calculations.

Table A14: Age Distribution

Age(Years)	Freq.	Percent	Cum.
5	571	24.02	24.02
6	582	24.48	48.51
7	589	24.78	73.29
8	413	17.37	90.66
9	222	9.34	100.00
Total	2,377	100.00	

Source: KiGGS Version 4, own calculations.

## H MTE

The MTE is given by:

$$MTE(x, p) = EY_1 - Y_0|X = x, P(Z) = u_D = \frac{dEY|X = x, P(Z) = p}{dp} \quad (5)$$

For the implementation the following equation has to be estimated

$$EY|X = x, P(Z) = p = \alpha_0 + X\beta_0 + (\alpha_1 - \alpha_0)p + X(\beta_1 - \beta_0)p + K(p) \quad (6)$$

where 0 indicates parameters of the equation for the non-treated outcome and 1 indicates parameters of the treated outcome. Here we use the parametric version to estimate the control function  $K(p)$ , assuming normally distributed error terms. Exact description of the estimation can be found in Heckman et al. (2006b) and Brave and Walstrum (2014). Normality is a strong assumption, but non-parametric methods are hard to justify in relatively small samples. This calculation is only performed to show that there is some evidence for treatment effect heterogeneity related to unobservable factors.

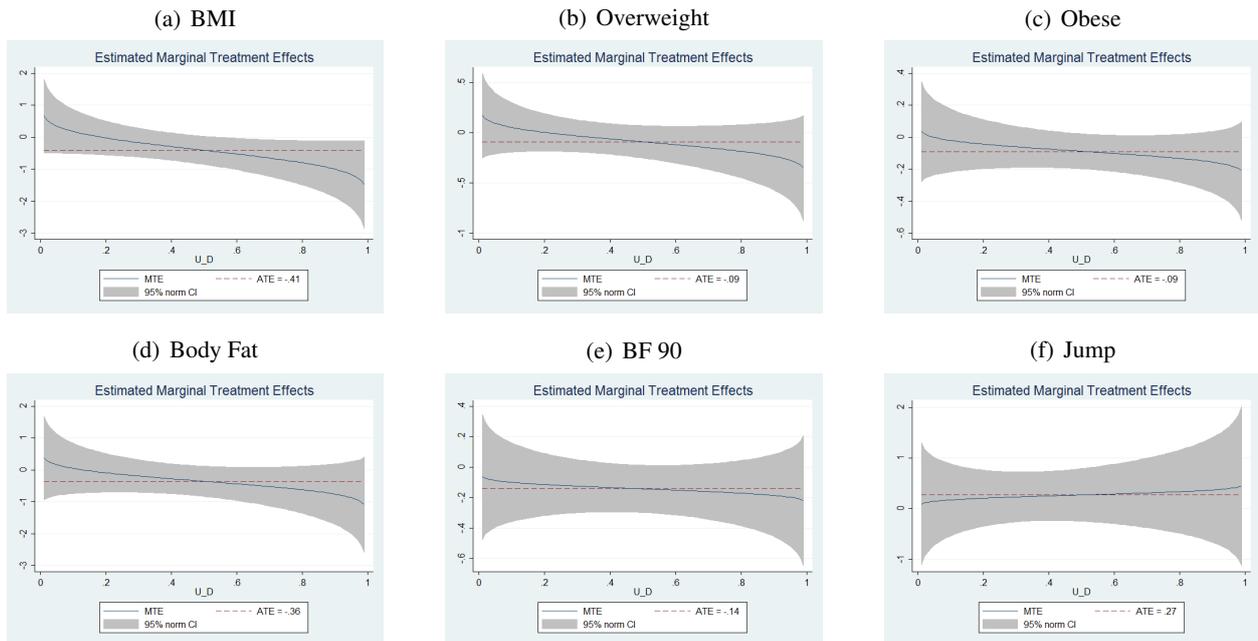


Figure A1: MTE distribution derived using the parametric normal approximation to the control function. *Notes:* Estimated controlling for the set of variables defined in ‘Child characteristic’, ‘Parents and Family’ and ‘Regional controls’ as described in table A1. *Source:* KiGGS Version 4 & Federal Statistical Office, own calculations.