Tjalling C. Koopmans Research Institute

Discussion Paper Series nr: 15-08

Quasi-experimental evidence on the relationbetween child care subsidies and child care quality

Yusuf Emre Akgunduz
Egbert Jongen
Paul Leseman
Janneke Plantenga

Universiteit Utrecht

Utrecht School of Economics

Tjalling C. Koopmans Research Institute
Utrecht School of Economics
Utrecht University

Kriekenpitplein 21-22
3584 EC Utrecht
The Netherlands
telephone +31 30 253 9800
fax +31 30 253 7373
website www.koopmansinstitute.uu.nl

The Tjalling C. Koopmans Institute is the research institute and research school of Utrecht School of Economics.
It was founded in 2003, and named after Professor Tjalling C. Koopmans, Dutch-born Nobel Prize laureate in economics of 1975.

In the discussion papers series the Koopmans Institute publishes results of ongoing research for early dissemination of research results, and to enhance discussion with colleagues.

Please send any comments and suggestions on the Koopmans institute, or this series to J.M.vanDort@uu.nl

ontwerp voorblad: WRIK Utrecht

How to reach the authors

Please direct all correspondence to the second author.
Janneke Plantenga
Utrecht University
Utrecht School of Economics
Kriekenpitplein 21-22
3584 TC Utrecht
The Netherlands.
E-mail: j.plantenga@uu.nl
Yusuf Emre Akgunduz
Egbert Jonge
The CPB Netherlands Bureau for Economic Policy Analysis
Van Stolkweg 14,
2585 JR, the Hague,
The Netherlands.
E-mail: Y.E.Akgunduz@uu.nl
E-mail: E.L.W.Jongen@cpb.nl
Paul Leseman
Department of Pedagogical and Education Sciences
Utrecht University
Martinus J. Langeveldgebouw
Heidelberglaan 1
Kamer F224
3584 CS Utrecht
The Netherlands
E-mail : P.P.M.Leseman@uu.nl

Utrecht School of Economics
Tjalling C. Koopmans Research Institute
Discussion Paper Series 15-08

Quasi-experimental evidence on the relationbetween child care subsidies and child care quality

Yusuf Emre Akgunduz[ab]
Egbert Jongen[b]
Paul Leseman[c]
Janneke Plantenga[a]

[a]Utrecht School of Economics
Utrecht University

[b]Centraal Plan Bureau
the Hague, the Netherlands

[c] Department of Pedagogical and Education Sciences,
Utrecht University

Update: June 2015, previously published as discussion paper 13-18

Abstract

In this article we present quasi-experimental evidence on the relation between child care subsidies and child care quality. We exploit the difference in funding of private and public centers in the Netherlands. A recent subsidy cut reduced funding for private centers while funding for public centers was unaffected. The quality measurements are from a panel survey in which centers' quality was evaluated through classroom assessments by trained observers. Using differences-in-differences we find that the subsidy reduction caused a statistically significant decrease in quality of one fifth of a standard deviation. We also present results for nonlinear differences-in-differences estimators. The decline in quality is robust across specifications and appears to be driven by a decline in the middle of the distribution. A limitation of our data set is that our pre-reform period is short, so that we cannot perform pre-reform placebo tests.

Keywords: child care subsidies, child care quality, DID, CIC

JEL classification: C21, I21, I28, J13

1 Introduction

The past decades have witnessed a strong rise in the enrollment of children in child care and preschool programs, following the rise in maternal employment and the increased availability and generosity of these programs. As a result, an increasing number of children are exposed to formal care with different levels of quality. Recent studies have shown that the quality of formal care is important for child development. Indeed, good quality care leads to more favorable outcomes later in life in terms of cognitive-academic achievements (Vandell et al., 2010), and employment, wage and crime rates (Heckman et al., 2010; Havnes and Mogstad, 2011). However, not all child care is of sufficiently good quality to raise child development, as evidenced by a number of studies that find a negative effect of child care attendance on child development (Baker et al., 2008; Datta Gupta and Simonsen, 2010; Herbst, 2013). Using a number of instruments, governments seek to promote good quality care. Indeed, governments set and enforce minimum standards for child-staff ratio's and the educational attainment of caregivers. However, these variables generally do not appear to be closely related to either the quality of care or child development (Blau, 1997, 1999, 2000). Another important policy that may influence formal care quality is subsidies to formal care, which may be an important resource for enhancing the quality of formal care. However, the extent to which subsidies to formal care affect the quality of formal care is largely unknown. Indeed, only a few, recent papers consider the effect of subsidies on care quality (Herbst and Tekin, 2010; Johnson et al., 2012). However, these studies rely on cross-sectional variation for identification, which raises concerns about whether or not they uncover a causal relation.

In this paper we use differences-in-differences (DID) to uncover the impact of child care subsidies on the quality of child care. Specifically, we study the effects of the 2012 subsidy cut for private child care centers in the Netherlands on the quality of these centers, using public child care centers, that were not affected by the reform, as a control group.

The two-tiered child care system in the Netherlands, where private and public child care centers coexist, allows for our quasi-experimental identification strategy. The treatment group consists of the private child care centers that were affected by the subsidy cuts from the central government. The control group consists of public child care centers, financed by municipalities, that were not affected by the subsidy cuts. We supplement the results of the linear DID model with estimates using quantile regression, changes-in-changes and recentered influence function models (Athey and Imbens, 2006; Firpo et al., 2009), which allow us to study heterogeneous treatment effects across the quality distribution. Our data come from the Pre-Cool survey, a two-wave, geographically-representative survey of child care centers in the Netherlands. The Pre-Cool survey has a panel structure, with most centers visited in both waves. The first wave of the survey was collected before the subsidy cut, the second wave was collected after the subsidy cut. The Pre-Cool survey focusses on the process quality of child care. Process quality can be broadly defined as measuring the quality of children's experiences in classrooms, particularly with regards to social and cognitive development (Blau, 2000). Process quality is measured through observations by trained personnel in centers' classrooms according to the Classroom Assessment and Scoring System (CLASS), a standardized assessment tool. This quality measure has been shown to be closely linked to child development.[1]

Our main findings are as follows. First, our baseline estimate for the average treatment effect shows that the cut in subsidies caused a statistically significant drop in average child care quality of one fifth of a standard deviation. Second, we find that the drop is primarily in the following elements of quality: i) regard for child perspectives, ii) behavior guidance, iii) facilitation of learning and development and iv) the quality of feedback. Third, the drop in quality appears to be driven by a decline in the middle of the quality distribution.

There are only a few empirical studies that analyze the link between child care subsidies and child care quality. Johnson et al. (2012) study the effect of subsidy receipt on child care quality for a cross-section of families with children 4 years of age in the US. Their treatment group consists of families that are eligible for child care subsidies and receive them. Their control group consists of families that are eligible for child care subsidies but do not receive them. Selection into subsidies is related to education and income levels of families, both because parents' own self-selection and administration agencies' targeting particular groups more agressively. They find that the process quality of child care is significantly higher for children of families that receive child care subsidies when compared to parents who do not receive child

[1] See Pianta et al. (2008) for a description of the CLASS methodology, and Pakarinen et al. (2010) for a recent validation.

care subsidies (and that do not participate in other public programs like Head Start), in the order of one third of a standard deviation. Herbst and Tekin (2010) study the relation between the receipt of child care subsidies and child development across counties in the US. They use data on cognitive, behavioral and physical development outcomes to measure the effect of receiving child care subsidies. They use county fixed effects as an instrument for the probability of receiving child care subsidies. They find that subsidy receipt results in a significant reduction of one quarter to one third of a standard deviation in math and reading scores. They argue that this negative effect may be due to families receiving child care subsidies purchasing lower quality care.

However, determining a causal effect of child care subsidies on child care quality using cross-sectional data is tricky. The descriptive statistics in Johnson et al. (2012) show substantial differences between the treatment and control group in e.g. the probability of the mother being in fulltime employment and household income. They do control for an extensive set of observable differences between treatment and control group, but this does not rule out unobservable differences between both groups that might also lead to a relation between the receipt of child care subsidies and child care quality. The instrument used by Herbst and Tekin (2010) captures differences in the receipt of child care subsidies across counties, but may also capture other differences across counties that may lead to a relation between the receipt of child care subsidies and child development beyond the county policy indicators they control for. We believe that the targeted reform in the Netherlands generates credible exogenous variation in child care subsidies. Combined with the panel structure of our data, so that we can control for unobserved fixed differences in child care quality between the treatment and control group, this allows for a more straightforward identification of the causal effect of child care subsidies on child care quality.

The outline of the article is as follows. Section 2 considers the institutional setup of child care in the Netherlands and the reform we use as exogenous variation. Section 3 outlines our empirical methodology. Section 4 discusses the data set used and gives some descriptive statistics. Section 5 presents the results and Section 6 concludes. An appendix contains supplementary material.

2 Institutional setting

The two-tiered structure of the Dutch child care sector is central to our identification strategy. Formal child care centers in the Netherlands can be divided into private daycare centers (*kinderdagverblijf*) and public playgroups (*peuterspeelzaal*). Daycare centers are typically used by dual-income families and working single parents and can cover enough hours for fulltime employment.[2] Playgroups provide 'parttime' care, between two and four half days per week, and are also used by families where one or more parents are not working. Both center types essentially provide the same child care services in terms of quality though (see the data section below).

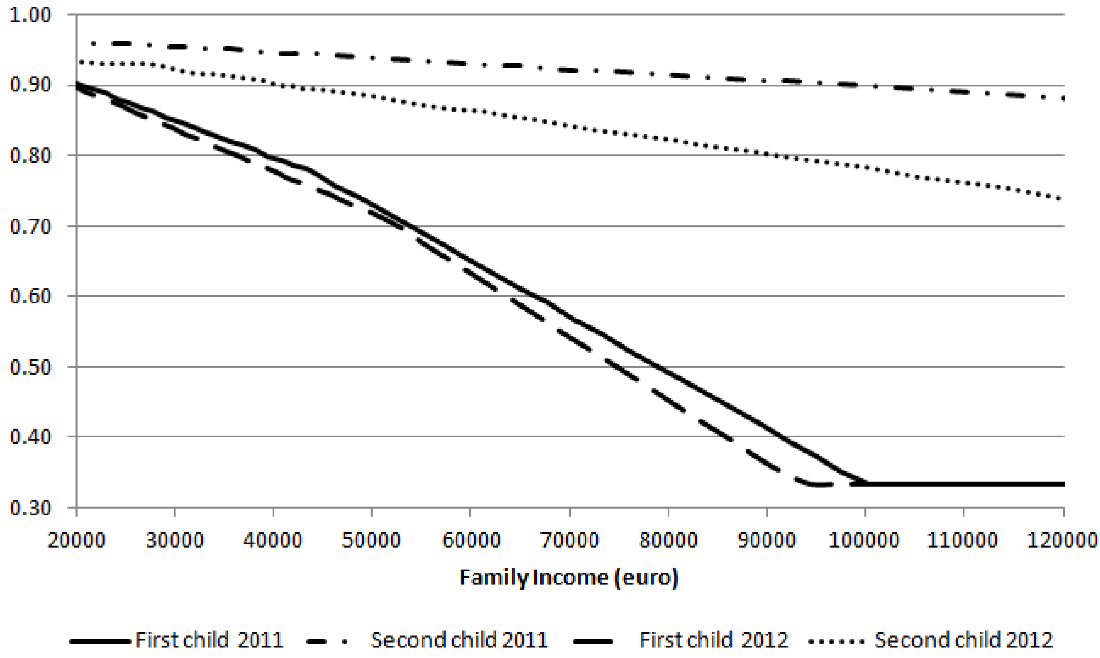
Daycare centers operate in a private market and parents are free to choose the daycare center they prefer.[3] Child care subsidies are paid to parents by the central government. Subsidies are paid per hour of care, up to a maximum price per hour beyond which parents receive no (additional) subsidy. The subsidy depends on income, with low incomes receiving a larger subsidy per hour than high incomes. The subsidy per child per hour also depends on the number of children in daycare per household; the subsidy is higher for the second (third etc.) child in daycare. In the period 2005–2008, there was a boost in the daycare sector, as subsidies for parents were increased substantially, cutting the effective parental fee for formal child care in half (Bettendorf et al., 2015). However, the reform was so succesful in terms of the use of formal child care that it increased public spending on formal child care from 1 billion euro in 2004 to 3 billion euro in 2009. This strong increase in public spending on child care, and the need for budget cuts following the financial crisis, led the government to curb public spending on child care.

In 2011 the government announced a substantial reduction in child care subsidies, which came into effect in 2012 (Ministry of Social Affairs and Employment, 2011). The change in subsidies in 2012 is illustrated in Figure 1. Subsidy rates were cut across the board for the first child by between 2 to 5 percentage points, while subsidy rates for the second (third etc.) child were reduced by more than 10 percentage

[2]In principle, daycare subsidies are paid only to working single parents and couples with children where both parents are working. However, parents that recently became unemployed remain eligible for daycare subsidies for a period of three months. Furthermore, parents participating in active labor market policies are also eligible for daycare subsidies.

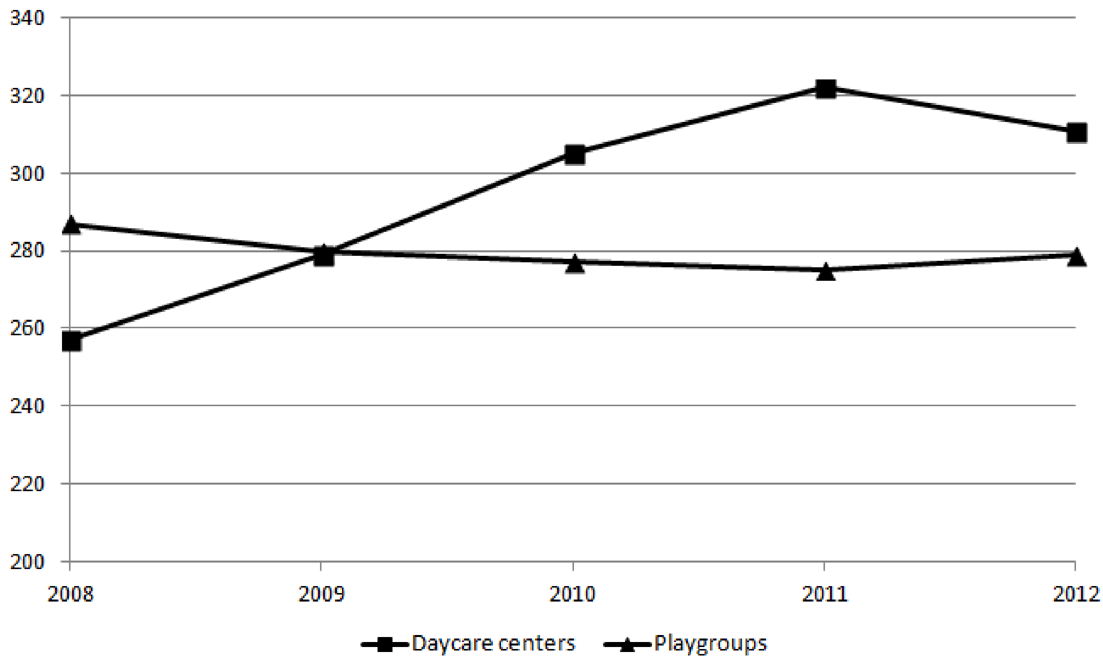
[3]Waiting lists are of limited importance in the Netherlands (Van Rens and Smit, 2011).

Figure 1: Subsidy rates for first and second child in daycare centers



Source: Tax Office.

Figure 2: Number of children in daycare centers and playgroups (x1000)



Source: Statistics Netherlands.

points depending on income.[4] Appendix A illustrates how these changes affected the net monthly expenditures on daycare for different family configurations (single parents and couples with young children), income levels and number of children in daycare. The increase in the parental fee was around 20, 40 and 50% for parents with one child, two children and three children in daycare, respectively. The increase in the parental fee was larger in absolute terms for middle and higher income families.

The subsidy reduction led to a negative demand shock for daycare. Indeed, 2012 is the first year since the introduction of the Law on Childcare (*Wet kinderopvang*) in 2005, in which hourly daycare prices rose less than the core inflation rate. Furthermore, after a steady rise up to 2011, the number of children in daycare declined by 3.5% from 2011 to 2012, see Figure 2. To remain solvent, the negative demand shock put pressure on daycare centers to cut costs. Our analysis shows that these cost savings led to a decrease in the process quality of child care. Parents may have been largely unaware of the reduction in quality, as the quality of child care is notoriously hard to determine by parents (Mocan, 2007).[5]

Playgroups were not affected by this subsidy cut. Playgroups are funded directly by municipalities, with minor parental contributions. Playgroups may have been affected by the recession in the Netherlands, but the same is true for daycare centers, regardless of the subsidy cuts. Furthermore, the austerity measures on the national level do not appear to have led to budget cuts for playgroups on the municipality level. This is supported by the fact that the number of pupils in playgroups has remained rather stable, see Figure 2, as opposed to the reversal of the growth of children in daycare centers. This supports our use of playgroups as the control group in our empirical analysis.

Both daycare centers and playgroups are regulated in terms of structural quality factors such as child-staff ratio's, educational attainment of caregivers and space

[4]In addition to the reduction in subsidies up to the maximum hourly price, the maximum hourly price was also kept constant (was not indexed with inflation) at €6.36 from 2011 to 2012. The average hourly price of formal child care rose from €6.32 in 2011 to €6.45 in 2012 (Ministry of Social Affairs and Employment, 2013). This further increased the effective costs of child care for parents from 2011 to 2012.

[5]A negative demand shock could potentially increase average quality, when the centers with the lowest quality drop out of the market. However, this mechanism seems less relevant in our case, the quality of daycare centers observed in wave 1 and not observed in wave 2 is similar to the quality of daycare centers observed in both waves, see Table A.2 in the appendix.

specifications. Municipalities further inspect daycare centers on whether they comply with the quality regulations agreed upon by parental organizations and daycare providers. Playgroups are inspected by municipalities to check if they comply with the quality regulations set by national law. In addition, both center types need to pass the inspection of the Municipality Health Service (*Gemeentelijke Gezondsheidsdienst*) on e.g. fire safety, for parents of children in daycare centers to be eligible for subsidies and for managers of playgroups to receive funding.

3 Empirical methodology

To estimate the average treatment effect (ATE) of the subsidy reduction on child care quality we use a linear differences-in-differences (DID) model (see e.g. Angrist and Pischke, 2008). The DID model relies on the difference in quality of daycare centers and playgroups before the reform and after the reform. By taking the difference of the difference, the DID model controls for fixed differences in quality between daycare centers and playgroups, and for common time effects in the quality in both center types.

We use the group (center type) dummy β_g to indicate to which type of care the classroom observation belongs, which is 0 for playgroups and 1 for daycare. β_t indicates the wave of the observation, which is 0 for the first wave and 1 for the second wave. The treatment dummy follows from the interaction of the group and the wave dummy $DID_{gt} = \beta_g * \beta_t$, which is 1 if the observation is from daycare and from the second wave, and 0 otherwise. We assume that quality y_{ijt} for classroom i of center j in period t is then determined by the following linear DID model:

$$y_{ijt} = \beta_0 + \beta_g + \beta_t + \delta DID_{gt} + \beta_m + \beta_j + \varepsilon_{ijt}, \quad (1)$$

where the parameter of interest is δ.[6] We add two sets of additional controls.[7] The β_m are time fixed effects for each month-year combination, to capture quality differences between different months in different years, assumed constant across center

[6]For notational convenience, we suppress a coefficient for the other dummies in the regression.

[7]We further test the robustness of the estimates for the inclusion of the number of children and the number of staff per group. However, we prefer to exclude these controls from the regression as they are likely to be endogenous.

types. The β_j are center fixed effects, which capture constant quality differences across centers.[8] ε_{ijt} is an error term whose structure is discussed below.

Next, we estimate models that take into account the shift in the full distribution of quality values. These nonlinear models relax the additivity assumption of the DID model. We first estimate a quantile differences-in-differences (QDID) model. In the QDID model the treatment effects are estimated for separate quantiles. The QDID allows us to estimate equation (1) for each quantile while including month and center fixed effects.

Subsequently, we implement the changes-in-changes (CIC) model of Athey and Imbens (2006). In the CIC model, the distribution of quality in the treatment and control groups are assumed to be generated by a 'production function' $Y = h^I(u, t)$, where u are unobserved characteristics at time t, and I is an indicator for treatment. The restriction on the production functions h for the treatment and control groups is that they are monotonic and increasing in u. The treatment effect for each quantile is determined in the following way. Take a particular quality level in the treatment group in the first wave Y_{10} (where 1 indicates the treatment group, and 0 indicates the first wave) and determine its associated quantile q. Next, find the quantile q' in the control group that has the same quality level $Y_{00} = Y_{10}$ in the first wave. Then, find the second wave value Y_{01} for the quantile q' in the control group. Finally, to obtain the treatment effect for quantile q in the treatment group, take the difference in quality $Y_{01} - Y_{00}$ for quantile q' in the control group, and subtract this from the difference in quality $Y_{11} - Y_{10}$ for quantile q in the treatment group. Indeed, Athey and Imbens (2006) show that the complete counterfactual distribution of the treatment group F_{Y^N} (where $I = N$ indicates no treatment) can be obtained using:

$$F_{Y^N}(y) = F_{Y_{10}}(F^{-1}_{Y_{00}}(F_{Y_{01}}(y))). \tag{2}$$

Finally, we apply the recentered influence function method of Firpo et al. (2009). They show that the unconditional quantile regression can be redefined as a linear regression where the dependent variable is the probability that an observed outcome is greater than a given level. We follow the application of the recentered influence function method by Havnes and Mogstad (2015) who use a threshold differences-in-differences (TDID) model. We can determine the effect of the subsidy reduction

[8] We can not use classroom fixed effects because we do not have a panel identifier for classrooms, only for centers.

$TDID_i$ on the possibility that the observed quality y_{ijt} is greater than the quality level y' by estimating (using ordinary least squares):

$$Pr(y_{ijt} > y') = \gamma_{0y'} + \gamma_{gy'} + \gamma_{ty'} + \delta_{y'} TDID_{gt} + \gamma_{my'} + c_{jy'} + \upsilon_{ijt}. \qquad (3)$$

As discussed in more detail in Havnes and Mogstad (2015), the advantage of the CIC and TDID model over the QDID model is that they take into account the full change in the distribution of outcomes for the treatment and control group. The QDID model 'simply' gives the differences-in-differences value using the same quantiles in the treatment and control group. The CIC model instead determines the treatment effect using as a counterfactual not the same quantile, but the quantile corresponding to the same level of initial quality (treatment and control units are comparable in terms of initial outcomes rather than the initial position in the cumulative distribution). The TDID model takes an approach related to the CIC model by first finding the corresponding quantile in the control group that has the same level as initial level of quality, but then looks at the change in the proportion of the samples before and after the treatment to calculate the treatment effect.

For each nonlinear model we can also calculate an average treatment effect by averaging the differences between the realized distribution of the treatment group and the counterfactual distribution. There can be noticeable difference between the average treatment effect from the linear and nonlinear DID models, dependening on the heterogeneity in the outcomes (Ropponen, 2011).

For the linear DID model we report standard errors clustered at the center-wave level, to allow for a correlation in the error term at the center-wave level (the level of treatment).[9] This gives us 260 clusters, which is deemed sufficient for accurate inference, based on the large-sample properties of the estimator.[10] For the nonlinear

[9]Failure to account for correlation in the error term across observations at the cluster level, within and between periods, can cause overrejection in DID estimates (Bertrand et al., 2004; Donald and Lang, 2007). Indeed, we find that the clustered standard errors are larger than the robust standard errors.

[10]Kézdi (2004) explores the small-sample properties of cluster-robust standard errors in fixed effects models using Monte Carlo methods. For a number of data generating processes he estimates the sample distribution of the standard error of the parameter of interest, based on 10,000 samples for different numbers of clusters (10, 50 and 500). He concludes that when the number of clusters is 50, the relative deviation between the asymptotic standard error and the estimated standard error never exceeds 5%. For 10 clusters, the error in some cases rises to 16%.

estimators we can not use clustered standard errors and report robust standard errors instead.

An important limitation of our data set is that we can not test the validity of common time effects, using e.g. placebo reform dummies in pre-reform years. We have observations of just two years in our data set, one year before the reform and one year after the reform. If there are differential trends in the quality of daycare centers and playgroups, our treatment dummy will be biased. What do we know about a potential trend in the quality of each center type? The Dutch Consortium for Childcare Research (NCKO) has been measuring the quality of daycare centers in the Netherlands using the ECERS-K methodology since 1995. In 1995 the average quality was 4.8 (on a scale from 1 (bad) to 7 (excellent)). Following the increase in places in daycare centers, quality dropped to 4.3 in 2001 and to 3.6 in 2005. After 2005, subsidies for daycare became much more generous, causing a rapid expansion. Quality dropped to a meagre 3.0 in 2008. Over the period 2008–2011, the growth in daycare places leveled off, and 2012 witnessed a decline following the subsidy cut. During this period, average quality as measured by the ECERS-K scale recovered to 4.3 in 2012 (no observations using ECERS-K were made between 2008 and 2012) (NCKO, 2013). Hence, if there was a trend in daycare quality during our data period, unrelated to the reform, it is likely to have been positive. Unfortunately, for playgroups there is no measurement of quality prior to our data period. However, since there were no significant changes in the enrollment or funding of playgroups in the years up to the reform, it is unlikely that there was a trend in the quality of playgroups. Below we estimate a negative treatment effect of the subsidy cut on the quality of daycare centers. The considerations above suggest that this is more likely to be an underestimate than an overestimate (in absolute terms) of the effect of the subsidy cut on the quality of daycare centers.

4 Data

The data on child care quality comes from the longitudinal Pre-Cool survey of Dutch child care centers, parents and children. The Pre-Cool survey consists currently of two waves, one collected in 2010 and 2011 and the other in 2012.[11] The primary

[11]Future waves of Pre-Cool will track the children in primary schools, there will be no new observations on the quality of child care in the survey.

Table 1: Quality indicators before and after subsidy reduction

	Daycare		Playgroups		DID w/o
	Before	After	Before	After	controls
Process quality indicators					
Overall quality	4.052	3.529	4.304	3.953	–0.172
	(0.863)	(0.675)	(0.910)	(0.680)	
- Emotional support	5.011	4.446	5.054	4.636	–0.147
	(0.855)	(0.736)	(0.865)	(0.699)	
- Instructional support	3.093	2.611	3.554	3.270	–0.198
	(1.072)	(0.800)	(1.145)	(0.824)	
Structural quality indicators					
Child-to-staff ratio	5.323	5.165	4.881	5.112	–0.389
	(2.092)	(2.129)	(2.124)	(2.224)	
Observations	392	355	466	411	
Number of centers	53	53	77	77	

Standard deviations in parentheses.

purpose of the Pre-Cool survey is to track the development of Dutch children.

The unit of analysis in our study is an individual group/classroom in daycare centers or playgroups. The Pre-Cool survey involved sending trained observers to measure the quality of classrooms. The average observation period was about 20 minutes for each classroom, and observations are made for multiple classrooms in each center. We use a balanced panel of 130 daycare centers and playgroups for which we have complete data for the analysis.[12] The centers used in the analysis are from 38 geographically-representative municipalities (see Figure A.1 in the appendix).

Quality is measured using the Classroom Assessment Scoring System (CLASS), developed by developmental psychologists. Measures like CLASS are said to mea-

[12]The full sample consists of 166 centers, of which 36 centers had to be dropped because of missing data. We increased the sample to 157 centers with an unbalanced panel and estimated the base regression without center and municipality fixed effects. The results show that the negative effects are more significant and larger (in absolute terms) by up to –0.05. The average quality indicators of centers that are observed only in the first wave are presented in Table A.2 in the appendix.

sure what is called ‘process quality’ or the quality of the child-caregiver interaction in the classroom. Multiple studies confirm that process quality measurements using CLASS are positively associated with development outcomes (Mashburn et al., 2008; Sabol et al., 2013). Furthermore, Sabol et al. (2013) find CLASS to be the strongest predictor of child learning outcomes compared to ECERS-R and structural indicators such as staff-to-child ratios. CLASS consists of two domains: instructional support and emotional support. Each domain is made up of several dimensions on which the observers grade the classroom interaction. The emotional support dimensions[13] are: i) positive climate, ii) teacher sensitivity, iii) behavior guidance, and iv) regard for child perspectives. Instructional support consists of: i) facilitation of learning and development, ii) quality of feedback, and iii) language development. Each dimension is graded by the observer on a discrete scale from 1 (low) to 7 (high). The average quality for each domain is constructed by simply taking the means of the dimensions, and the overall quality score is the mean of the two domain scores.[14]

Table A.2 presents descriptive statistics for the quality indicators for daycare centers and playgroups in wave 1 (‘Before’) and wave 2 (‘After’). The wave 1 observations were made in late 2010 and early 2011.[15] The wave 2 observations were made in 2012.[16] The average score for overall quality in the pre-reform period is around 4 for both daycare centers and for playgroups. For both child care types scores are higher for the emotional support domain than for the instructional support domain.

[13] An additional dimension named ‘Negative climate’ is collected but has low variation, is coded in reverse (1 being highest as opposed to 7) and its effect should be captured by the positive climate measure. Pakarinen et al. (2010) found a similarly low variation in Finland for the negative climate dimension and concluded that CLASS measures are a better predictor of classroom quality when negative climate is excluded. For the Pre-Cool survey, Leseman and Slot (2014) report a low variation for the negative climate measure, since almost all classrooms score very low on it.

[14] In development psychology factor analysis is sometimes used to construct quality scores, but we use the means in our regressions because they are easier to interpret. As a robustness check we used factor analysis to construct average domain scores. The factor loadings appear to be around 0.7 for all dimensions except the ‘Regard for child perspectives’ dimension which has a factor loading of 0.34. The estimated treatment effects when domain factor scores are used are similar to the effects on domain averages, with a larger effect for the instructional support domain compared to the emotional support domain. Further details available on request.

[15] Excluding the 13 centers whose wave 1 observations were made in 2010 does not affect the estimated average treatment effect, but reduces the statistical power of the regressions.

[16] In total, we have observations from 15 different months over the years 2010–2012, for which we include time fixed effects in the regressions.

Figure 3: Distribution of overall quality in daycare centers

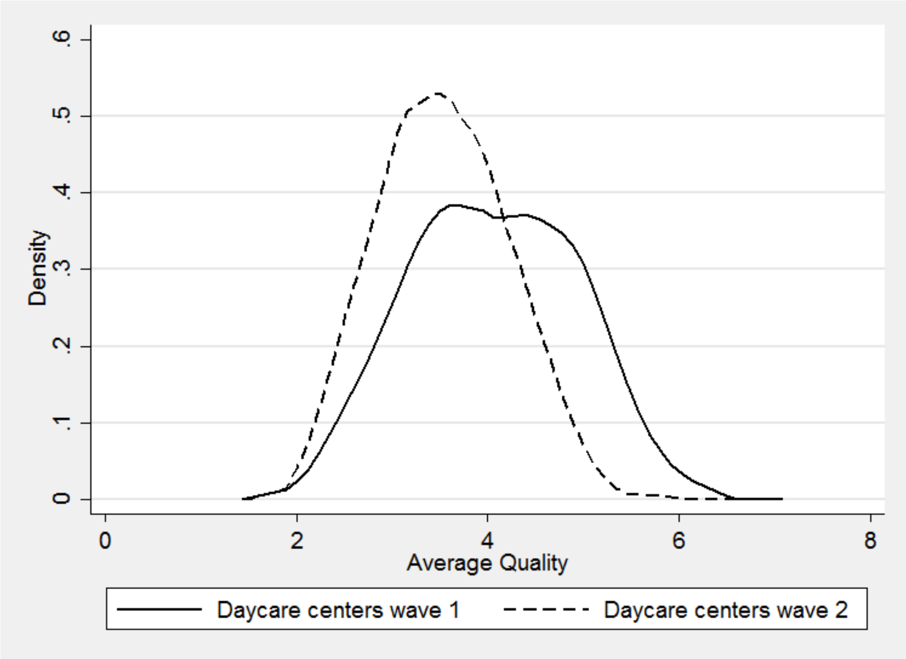
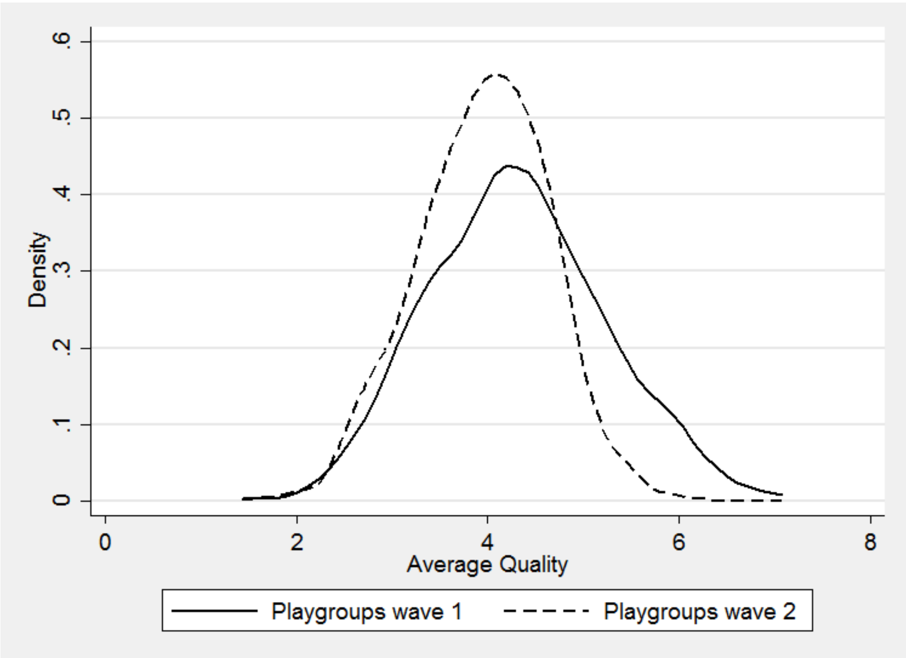


Figure 4: Distribution of overall quality in playgroups



Post-treatment scores are lower for both daycare centers and playgroups, but the drop in scores is more pronounced for daycare centers. The drop is most visible in the instructional support domain. A simple differences-in-differences calculation, using only group averages for the treatment and control group before and after the reform, suggests an average treatment effect of –0.172 for overall quality. The effect appears somewhat lower for emotional support (–0.147) and somewhat higher for instructional support (–0.198).

Looking at the structural quality indicator, we see a small drop in the child-to-staff ratio for daycare centers, and a small rise in the child-to-staff ratio for playgroups. A simple DID calculation suggests a drop in the child-to-staff ratio in daycare centers relative to playgroups by –0.389 points.

Next to the means we are also interested in changes in the full distribution of quality in daycare centers and playgroups. Figure 3 and 4 show a kernel density estimate of the distribution of the quality scores before and after the reform per child care type. In wave 1, the quality score distribution is quite similar for playgroups and daycare centers, though the top of the quality distribution for daycare centers appears somewhat wider. In wave 2, the quality distribution of daycare centers shifts noticeably more to the left than the quality distribution of playgroups. Furthermore, the top of the quality distribution for daycare centers becomes more 'peaked' after the reform.

5 Results

5.1 Average treatment effect

Table 2 gives the average treatment effect on overall quality, for different sets of controls. Column (1) gives the treatment effect using only a group dummy and one time dummy (for the wave). This model gives a drop in overall quality of –0.172 points. In column (2) we replace the single time dummy by year-month fixed effects to control for the timing of the observation over the year. The treatment effect drops to –0.194. In column (3) we add municipality fixed effects in order to control for heterogeneity in local child care markets, which causes a further drop in the treatment effect to –0.222. Column (4) then adds fixed effects for municipality interacted with center type (daycare/playgroup), to control for heterogeneity in local

Table 2: Average treatment effect on overall quality

Model	(1)	(2)	(3)	(4)	(5)
Overall quality	–0.1725	–0.1942	–0.2216	–0.2242	–0.2124
	(0.1424)	(0.1402)	(0.1313)*	(0.1193)*	(0.0953)**
Time controls	Wave	Year-month	Year-month	Year-month	Year-month
Municipality FE	No	No	Yes	No	No
Munic. x center type FE	No	No	No	Yes	No
Center FE	No	No	No	No	Yes
Observations	1,624	1,624	1,624	1,624	1,624
Number of centers	130	130	130	130	130

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by center-wave in parentheses (260 clusters).

markets by center type. The results are similar to column (3). Finally, column (5) adds center fixed effects for each individuals center. The treatment effect of –0.212 is similar to the models in column (3) and (4). The model in column (5) is our preferred specification because it is the most flexible in terms of the distribution of the unobserved fixed effects.[17]

The negative treatment effect on overall quality of around 0.2 is one third of a standard deviation. For the US, Duncan (2003) finds that a reduction of one standard deviation in process quality is associated with a decrease of one tenth of a standard deviation in cognitive scores for small children. If the relation between child care quality and child development in the Netherlands is similar to the US, the subsidy reduction would have decreased cognitive development of Dutch children attending daycare by about 3% of a standard deviation.

Table 3 shows which elements of quality drive the drop in quality. The individual elements are grouped into elements of the emotional support domain and elements of the instructional support domain. The coefficients are negative for all elements. In the emotional support domain, the drop is larger and more significant in the items 'Regard for child perspective' and 'Behavior guidance'. In the instructional support domain, the effect is larger for 'Quality of feedback' and in particular for 'Facilitation of learning and development'. Previous studies have shown that instructional

[17]When we include the child-to-staff ratio as an additional control variable in the model of column (5) the treatment effect is quite similar (–0.216).

Table 3: Average treatment effect on single elements of quality

	Average treatment effect
Overall quality	–0.2124
	(0.0953)**
- Emotional support	–0.1743
	(0.0924)*
- - Positive climate	–0.0549
	(0.1151)
- - Teacher sensitivity	–0.1244
	(0.1170)
- - Regard for child perspectives	–0.2569
	(0.1157)**
- - Behavior guidance	–0.2609
	(0.1152)**
- Instructional support	–0.2505
	(0.1177)**
- - Facilitation of learning and development	–0.4509
	(0.1301)***
- - Quality of feedback	–0.2348
	(0.1212)*
- - Language development	–0.0660
	(0.1471)

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by center-wave in parentheses. All regressions include year-month and center fixed effects.

Table 4: Average treatment effect structural quality indicators

	Average treatment effect
Number of children	–0.9734
	(0.4051)**
Number of adults	–0.1247
	(0.0923)
Child-to-staff ratio	–0.2182
	(0.2398)

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by center-wave in parentheses. All regressions include year-month and center fixed effects. Data is from 130 centers, as in the process quality regressions. However, in the structural quality regressions 10 classroom observations had to be dropped due to missing data on the number of children and/or the number of staff.

support measures are particularly important for school readiness (Mashburn et al., 2008). Noting that the scores for instructional support were already lower to begin with, a further decrease seems more likely to have negative consequences on child development.

Table 4 gives the average treatment effect on the structural quality indicator in our data set, the child-to-staff ratio. The drop in demand for daycare led to a drop in the child-to-staff ratio in daycare centers relative to playgroups. One might have expected that this would cause an increase rather than a decrease in child care quality. However, studies have shown that structural quality indicators are not a very good predictor of the quality of care or child development (Blau, 1997, 1999, 2000). Hence, we prefer to look at the treatment effects on the indicators of process quality considered above.

5.2 Heterogeneous treatment effects

Next, we consider which part of the quality distribution was most affected by the reform. Figure 3 and 4 reveal important differences in the shift of the quality distribution of daycare centers relative to playgroups. Table 5 gives the quantile treatment effects at the 10th, 25th, 50th, 75th and 90th quantile, along with the treatment effect averaged over all quantiles. A plot of the quantile effects for the

Table 5: Treatment effects by quality quantiles

	ATE	Quantile treatment effect				
Quantile		0.10	0.25	0.50	0.75	0.90
QDID						
Treatment effect	-0.1942	–0.2500	–0.2083	–0.2500	–0.3333	–0.1250
Robust s.e.	(0.0786)**	(0.1192)**	(0.0944)**	(0.0973)**	(0.1022)***	(0.1440))
CIC						
Treatment effect	–0.2217	–0.1250	–0.2917	–0.2500	–0.3333	–0.1667
Bootstrapped s.e.	(0.0660)***	(0.1063)	(0.1169)**	(0.1065)**	(0.0957)***	(0.0850)**
TDID						
Treatment effect	–0.2124	–0.1786	–0.4715	–0.3522	–0.2211	0.0471
Robust s.e.	(0.0712)***	(0.1307)	(0.1031)***	(0.1035)***	(0.1076)**	(0.1365)

*** p<0.01, ** p<0.05, * p<0.1. We use 200 repetitions for the bootstrapped standard errors in the CIC model. The QDID includes year-month fixed effects. The TDID model includes year-month fixed effects and center fixed effects.

different nonlinear estimators is given in Figure 5, 6 and 7.

The average treatment effect of the nonlinear models is very similar to the estimates from the linear model. However, there appear to be substantial differences by quantiles, though the treatment effects by quantile are typically not significantly different from the average treatment effect. The treatment effects by quantiles suggest that the effects are larger in the middle of the quality distribution, especially when using the CIC and TDID models. The QDID model suggests that the effects are strong also at the bottom of the quality distribution, but this is comparing treatment and control units by initial quantile, not by initial quality level.

A possible explanation for the relatively small drop in quality at the bottom of the quality distribution may be that these child care centers are used mostly by low income families. The increase in the parental fee was targeted more at middle and high income families than at low income families. Another possible explanation for the relatively small drop in quality at the bottom of the quality distribution may be that low quality child care centers simply do not have any more room to cut costs and lower quality given the current regulations.

The small drop in quality at the top of the distribution may reflect that high

Figure 5: Quantile effects - QDID

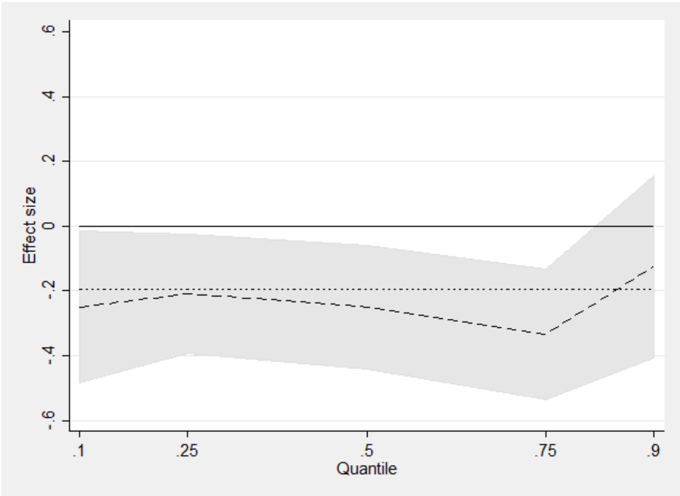


Figure 6: Quantile effects - CIC

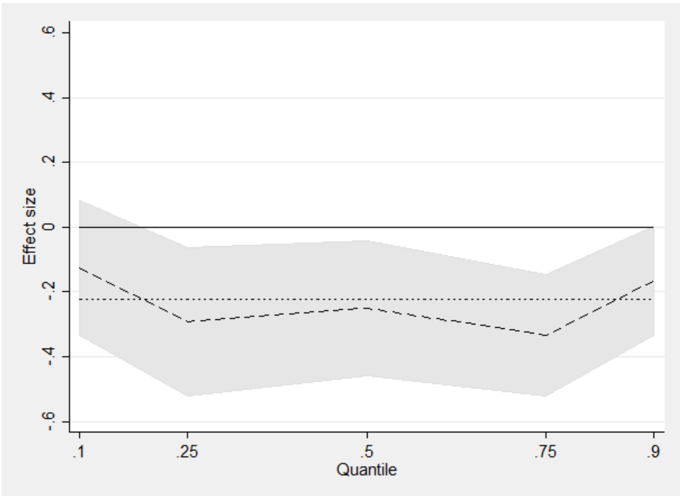
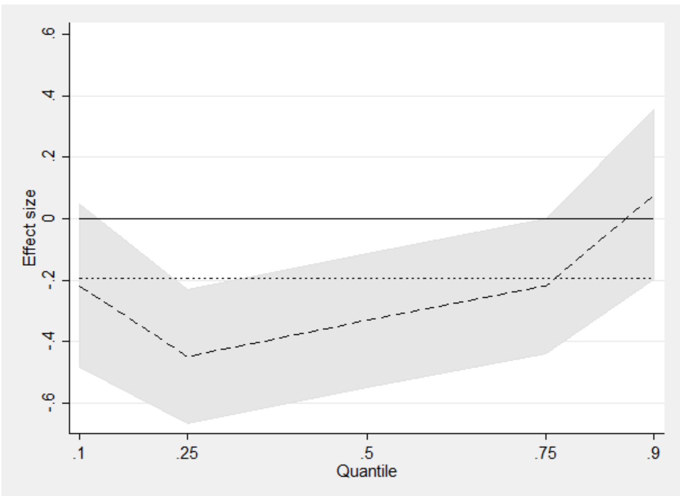


Figure 7: Quantile effects - TDID



quality centers are catering to parents with strong tastes for high quality child care, and these parents perhaps respond more strongly to a drop in child care quality than a rise in the effective hourly price.

6 Conclusions

The quality of child care is an important element of child care policy design. However, the impact of child care subsidies on child care quality is largely unknown. We present quasi-experimental evidence on the relation between child care subsidies and child care quality, using a recent reform in the Netherlands. The two-tiered child care system in the Netherlands, where private daycare centers (treatment group) and public playgroups (control group) coexist, allows for an identification of the causal effect using differences-in-differences. The results show that the quality in Dutch daycare centers declined as a result of the subsidy cut, by about one third of a standard deviation. Prior trends in daycare quality suggests that our results might even be underestimating the decline in quality caused by the subsidy reduction. We also consider the heterogeneity in the treatment effect using nonlinear models, our findings suggest that the decline in quality is strongest in the middle of the quality distribution, though the effects are typically not statistically different from the average treatment effect.

In the wake of the financial crisis, fiscal consolidation and austerity have led to cuts in public spending on child care and related programs. In the US, the 2013 budget sequestration resulted in cuts in the budget for the Head Start Program. In the Netherlands, there was a substantial reduction in subsidies for parents using formal child care in 2012. Many other European countries are also engaged in austerity measures to lower the public debt, and spending on child care and related programs is unlikely to be immune to these cuts in these countries as well. Our paper shows that for a complete picture or the pros and cons of cuts in public spending on these programs, policymakers should consider not only the effects on e.g. maternal employment, but should also consider the effects on the quality of care, as the quality of care is an important determinant of child development.

Acknowledgements

The authors gratefully acknowledge funding by the Bernard van Leer foundation and the CARE project funded by the European Union grant no 613318. We have received helpful comments and suggestions from Marcel van Aken, Leon Bettendorf, Jeroen Hinloopen, Thomas van Huizen, Daniel Possenriede, Dinand Webbink, Bas ter Weel, and seminar and congress participants at Utrecht School of Economics, CPB Netherlands Bureau for Economic Policy Analysis and ESPE 2014. Remaining errors are our own.

References

Angrist, J. and Pischke, J.-S. (2008). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press, Princeton.

Athey, S. and Imbens, G. (2006). Identification and inference in nonlinear difference-in-differences models. *Econometrica*, 74(2):431–497.

Baker, M., Gruber, J., and Milligan, K. (2008). Universal child care, maternal labor supply, and family well-being. *Journal of Political Economy*, 116(4):709–745.

Bertrand, M., Duflo, E., and Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics*, 119(1):249–275.

Bettendorf, L., Jongen, E., and Muller, P. (2015). Childcare subsidies and labour supply: evidence from a large Dutch reform. *Labour Economics*, forthcoming.

Blau, D. (1997). The production of quality in child care centers. *Journal of Human Resources*, 32(2):354–387.

Blau, D. (1999). The effect of child care characteristics on child development. *Journal of Human Resources*, 34(4):786–822.

Blau, D. (2000). The production of quality in child-care centers: another look. *Applied Developmental Science*, 4(3):136–148.

Datta Gupta, N. and Simonsen, M. (2010). Non-cognitive child outcomes and universal high quality child care. *Journal of Public Economics*, 94(1):30–43.

Donald, S. and Lang, K. (2007). Inference with difference-in-differences and other panel data. *Review of Economics and Statistics*, 89(2):221–233.

Duncan, G. (2003). Modeling the impacts of child care quality on children's preschool cognitive development. *Child Development*, 74(5):1454–1475.

Firpo, S., Fortin, N., and Lemieux, T. (2009). Unconditional quantile regressions. *Econometrica*, 77(3):953–973.

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

Havnes, T. and Mogstad, M. (2015). Is universal child care leveling the playing field? *Journal of Public Economics*, forthcoming.

Heckman, J., Moon, S., Pinto, R., Savelyev, P., and Yavitz, A. (2010). The rate of return to the HighScope Perry Preschool Program. *Journal of Public Economics*, 94(1-2):114 – 128.

Herbst, C. (2013). The impact of non-parental child care on child development: evidence from the summer participation "dip". *Journal of Public Economics*, 105:86–105.

Herbst, C. and Tekin, E. (2010). Child care subsidies and child development. *Economics of Education Review*, 29(4):618–638.

Johnson, A., Ryan, R., and Brooks-Gunn, J. (2012). Child-care subsidies: do they impact the quality of care children experience? *Child Development*, 83(4):1444–1461.

Kézdi, G. (2004). Robust standard error estimation in fixed-effects panel models. *Hungarian Statistical Review*, 9:96–116.

Leseman, P. and Slot, P. (2014). *Kwaliteit en curriculum van voorschoolse opvang en educatie in Nederland*. Nationaal Regieorgaan Onderwijsonderzoek, The Hague.

Mashburn, A., Pianta, R., Hamre, B., Downer, J., Barbarin, O., Bryant, D., Burchinal, M., Early, D., and Howes, C. (2008). Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. *Child Development*, 79(3):732–749.

Ministry of Social Affairs and Employment (2011). Vereenvoudiging en beperking kindregelingen. The Hague.

Ministry of Social Affairs and Employment (2013). Brief voor de Tweede Kamer: Cijfers kinderopvang 2012. Technical report, Ministry of Social Affairs and Employment, The Hague.

Mocan, N. (2007). Can consumers detect lemons? An empirical analysis of information asymmetry in the market for child care. *Journal of Population Economics*, 20:743–780.

NCKO (2013). Pedagogische kwaliteit van de kinderopvang voor 0- tot 4-jarigen in Nederlandse kinderdagverblijven in 2012. Technical report, Nederlands Consortium Kinderopvang Onderzoek.

Pakarinen, E., Lerkkanen, M.-K., Poikkeus, A.-M., Kiuru, N., Siekkinen, M., Rasku-Puttonen, H., and Nurmi, J.-E. (2010). A validation of the classroom assessment scoring system in Finnish kindergartens. *Early Education & Development*, 21(1):95–124.

Pianta, R., LaParo, K., and Hamre, B. (2008). *The Classroom Assessment Scoring System Pre-K Manual*. Baltimore, MD: Brookes.

Ropponen, O. (2011). Reconciling the evidence of Card and Krueger (1994) and Neumark and Wascher (2000). *Journal of Applied Econometrics*, 26(6):1051–1057.

Sabol, T. J., Soliday Hong, S., Pianta, R. C., and Burchinal, M. R. (2013). Can rating pre-K programs predict childrens learning. *Science*, 341(6148):845–846.

Van Rens, C. and Smit, F. (2011). Wachtlijsten en wachttijden kinderdagverblijven en buitenschoolse opvang - 6e meting. ITS, Radboud University, Nijmegen.

Vandell, D. L., Belsky, J., Burchinal, M., Steinberg, L., and Vandergrift, N. (2010). Do effects of early child care extend to age 15 years? Results from the NICHD study of early child care and youth development. *Child Development*, 81(3):737–756.

Appendix

Change in child care costs for families

Table A.1 shows the increase in child care costs for families with one, two or three children in daycare. We calculate child care costs for single parents and couples using three days of daycare per week. For single parents we calculate child care costs at median income. For couples we calculate child care costs at 150% and 200% of median income. The table shows that child care costs have increased by by 26% for single parent families with one child and the increase is 46% for two and 58% for three children. Child care costs rose more for higher income families, even though the percent change is lower due to the higher base cost in 2011. For couples with 150% of the median income, the percent change is between 23% and 52% depending on the number of children. The increases are similar for couples with 200% of the median income with an increase if 21% for parents with one child in daycare and 52% for those with three.

Table A.1: Child care costs of households in 2011 and 2012 (in euro)

	2011		2012		Percentage
	Income	Net cost	Income	Net cost	change in net cost
Single parent, median income					
One child	32,500	120	33,150	152	+26
Two children	32,500	155	33,150	226	+46
Three children	32,500	190	49,725	301	+58
Couple, 1.5x median income					
One child	48,750	183	49,725	225	+23
Two children	48,750	227	49,725	323	+42
Three children	48,750	271	49,725	422	+52
Couple, 2x median income					
One child	65,000	278	66,300	336	+21
Two children	65,000	332	66,300	461	+39
Three children	65,000	385	66,300	586	+52

Figure A.1: Number of class observations per municipality

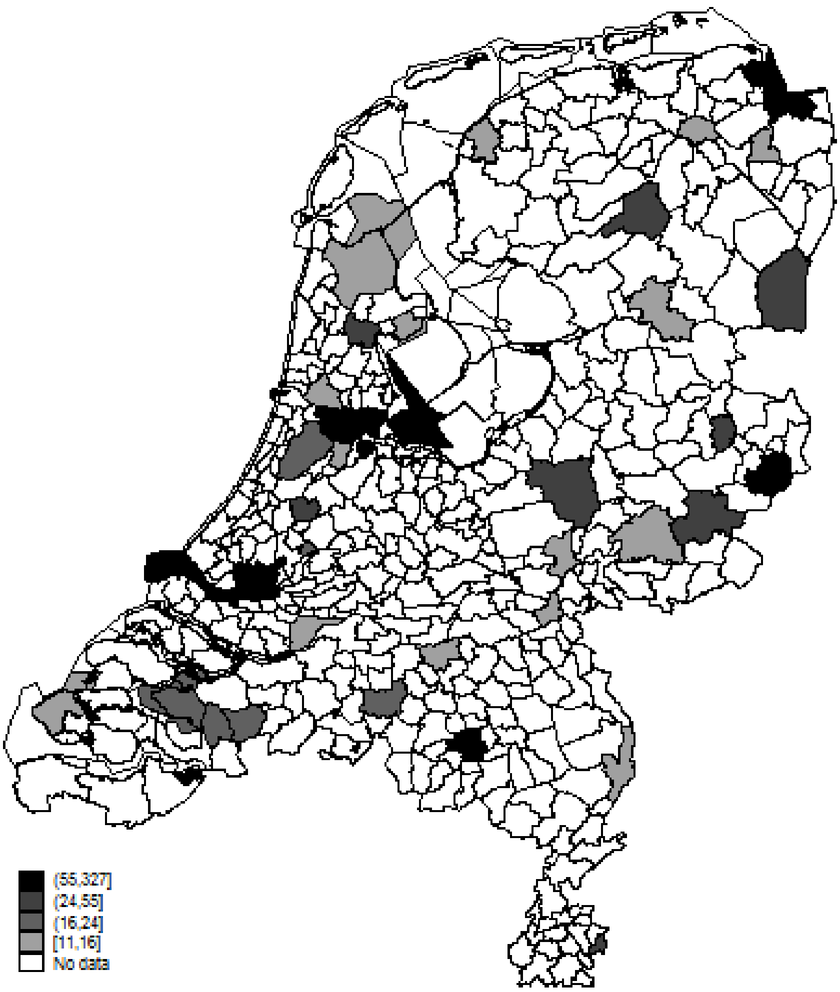


Table A.2: Summary statistics for centers missing in wave 2

	In both waves		First wave only	
	Daycare	Playgroup	Daycare	Playgroup
Process quality indicators				
Overall quality	4.052	4.304	4.017	3.957
	(0.863)	(0.910)	(0.657)	(0.830)
- Emotional support	5.011	5.054	5.048	4.874
	(0.855)	(0.865)	(0.682)	(0.919)
- Instructional support	3.093	3.554	2.985	3.040
	(1.072)	(1.145)	(0.904)	(0.935)
Structural quality indicators				
Child-to-staff ratio	5.323	5.165	5.094	4.990
	(2.092)	(2.129)	(1.855)	(2.021)
Observations	392	466	88	92
# of centers	53	77	11	16

Standard deviations in parentheses.