The effect of time spent in kindergarten on student achievement

By

TÍMEA LAURA MOLNÁR

Submitted to

Central European University

Department of Economics

In partial fulfillment of the requirements for the degree of Master of Arts

Supervisor: Professor Gábor Kézdi

Budapest, Hungary

2011

Abstract

I address the causal effect of time spent in kindergarten on student achievement. I present a theoretical model of the parental decision of how many years the child should spend in kindergarten. I derive the probability limit of the Ordinary Least Squares (OLS) and the Instrumental Variable (IV) estimators. Using Hungarian data from 2008 for 6th graders, I estimate OLS and IV estimations. I use scarcity of kindergarten teachers in the municipality the child lived at her age of 4 as the instrument. The OLS results show that an additional year spent in kindergarten increases student achievement by 6% standard deviation in mathematics and reading (controlling for social, compositional and spatial characteristics of the family and the school-starting age of the child). The IV results show that, keeping the same controls fixed, an additional year spent in kindergarten increases student achievement by 18.4% standard deviation in mathematics and 21.5% standard deviation in reading. If appropriate assumptions are satisfied then the OLS and IV estimates are lower and upper bounds of the population effect, respectively. The effects are systematically stronger for disadvantaged children. My results suggest that less privileged families are more constrained in case of scarcity of kindergarten service provision, which raises inequality concerns.

Acknowledgements

First of all, I would like to express my gratitude to Professor Gábor Kézdi, not only for his helpful comments and continuous supervision, but for the inspiring and encouraging atmosphere he provided during our consultations. I am especially lucky for having had the opportunity to take three of his courses at CEU, and to work on my thesis under his close supervision.

My thanks go to Gábor Kertesi for providing me the municipality-level demography dataset for Hungary.

I would also like to thank my CEU classmates, Ildikó Magyari and Jenő Pál for their help with the theoretical part of my thesis and for their continuous encouragement throughout the year.

Thomas Rooney, my Academic Writing Instructor provided invaluable help in proof-reading my thesis.

Last, but in no way least, let me take this opportunity to thank my parents and my sister for their patience and support.

Table of Contents

List of Tables

Appendix

List of Figures

Chapter 1: Introduction

In this thesis I address the causal effect of time spent in kindergarten on student achievement. I present a theoretical model of the parental choice of the length of kindergarten attendance and derive the probability limit of the Ordinary Least Squares (OLS) and the Instrumental Variable (IV) estimator. I show that the IV estimate is likely to be upward biased. The direction of the bias in the OLS estimate is more ambiguous, but it is slightly more likely to be downward biased if particular assumptions are satisfied. If these are satisfied, I provide a lower and an upper bound for the population effect. I present OLS and IV estimates of time spent in kindergarten on mathematics and reading standardized test scores, by making use of Hungarian data from 2008 for 6th graders. I use scarcity of kindergarten teachers in the municipality the child lived at her age of 4 as the instrument.

According to the OLS results, an additional year spent in kindergarten increases student achievement by 6% standard deviation in mathematics and reading (controlling for social, compositional and spatial characteristics of the family and the school-starting age of the child). According to the IV results, keeping the same controls fixed, an additional year spent in kindergarten increases student achievement by 18.4% standard deviation in mathematics and 21.5% standard deviation in reading. I find systematically stronger effects for disadvantaged children. There is also a clear pattern of less privileged families being more constrained in case of scarcity of kindergarten service provision.

The causal effect of time spent in kindergarten on student achievement is particularly important if one seeks the opportunities to reduce disparities in an educational system. There is sound empirical evidence that some disadvantaged children enter school with such substantial disparities that the school system cannot remedy the differences in student achievements later (see e.g. Brooks-Gunn, 2003).

Kindergarten care contributes to early childhood development. Spending more time in kindergarten could, at least partially, outweigh the insufficient positive stimuli received at home and help disadvantaged students to start school with relatively fewer disparities. If kindergarten attendance has a positive effect on less privileged students' achievement, then this indicates that kindergarten care as a form of early childhood intervention is effective and can serve as a tool to reduce educational disparities.

Besides equity concerns, the question is also important from efficiency point of view. According to evidence in life cycle skill formation, remediation of inadequate early human capital investments is very costly at later stages of the life cycle. At the same time, there is no equity-efficiency trade-off in human capital investments at early stages (Heckman and Carneiro, 2003). Consequently, if kindergarten attendance has a positive effect on student achievement, then it is not only an effective tool to reduce educational disparities, but is also relatively cheap one.

The importance of early childhood development has been widely stressed not only in psychology, neuroscience and cognition (e.g. Shonkoff, 2003) but recently also in economics (see, among others, Heckman and Carneiro (2003) or Currie (2001)). Cunha and Heckman (2007) provide a model of life cycle skill formation that explains what kind of technological reasons are behind the observation that ability gaps between disadvantaged and affluent children are present already at early ages. The authors emphasize the dynamic nature of life cycle skill formation in which self-productivity and dynamic complementarity are crucial. Self-productivity means that skills acquired at one stage in the life cycle augment the skills that are acquired at later stages. Thus, skills are "self-reinforcing and cross-fertilizing" (Cunha and Heckman, 2007, pp. 8.). According to dynamic complementarity, investment at one stage facilitates the productivity of human capital investment at later stages (Cunha and Heckman, 2007).

Primarily due to the aforementioned characteristics of life cycle skill formation, gaps can be reduced if intervention is present already at early stages in the life cycle. At the same time, correction of inadequate early investments is costly at later stages and yield higher returns for the more able (Cunha and Heckman, 2007). Since disadvantaged children are less likely to receive substantial early investment, investment in disadvantaged young children has been proven to be the most efficient (Cunha et al. (2006) present a comprehensive survey of this empirical evidence). The positive long and short-run effect of special pre-school care programs targeted on disadvantaged children is also well-documented (e.g. Blau and Currie (2006) provide a comprehensive survey, Ludwig and Miller (2007) show evidence about Head Start, Heckman et al. (2010) about Perry Preschool Program and Chetty et al. (2011) about Project STAR).

On the contrary, there is little empirical evidence available on the benefits of publicly provided general kindergarten care. Berlinski et al. (2009) estimate the effect of expansion of public kindergarten education on student achievement in Argentina. Since the authors cannot observe both kindergarten attendance and test scores for the same individuals, they measure the intention-to-treat effect of constructing an extra kindergarten place per 3-5 aged child on 3rd grade test score. They find that the impact of an additional place on 3rd grade test score is 8%, *ceteris paribus*. They also find that the children' classroom attention, effort, discipline, and participation are positively affected by kindergarten attendance. Magnuson et al. (2007) find that kindergarten attendance enhances mathematics and reading skills at school entry, but the effects disappear by the spring of the first grade. They find the largest and most lasting benefits for disadvantaged children. Neither of the analysis of Berlinski et al. (2009) or Magnuson et al. (2007) provide theoretical foundations for the choice of kindergarten attendance. It remains unclear through which channels this choice is affected by, among others, availability of kindergarten services, potential benefits or socio-economic status. Without theoretical foundations one cannot assess the relationship between the obtained OLS and IV estimates.

The contribution of my thesis is both theoretical and empirical.

From the theoretical point of view, I present a model of parental choice of the length of the child's kindergarten attendance. This choice is exactly the mirror image of the standard problem of how many years to spend in school. The effect of time spent in kindergarten on student achievement is a form of returns to schooling. Analogously to the standard literature I assume that children are heterogeneous in their abilities acquired at home and in their benefits from kindergarten attendance. Additionally, I assume that families are heterogeneous in their costs of the child's kindergarten attendance. This model provides a framework to derive the probability limit of the OLS and IV estimators and to understand the role of individual-specific components in the (potential) bias of the estimators.

From the empirical point of view my estimates add to the (smaller) literature of the effect of general kindergarten care on student achievement. This issue is of particular importance in Hungary now, since the ministry responsible for public education contemplates to extend the length of compulsory kindergarten attendance for 1 year to 3 years[1]. Thus, my results can be informative also for policy-makers. They imply that the extension could lead to substantial positive effects for disadvantaged children. At the same time, wider and cheaper access to kindergarten services for less privileged families should be facilitated, since exactly these families seem to be more constrained in case of scarcity.

The thesis is organized as follows. In Chapter 2 I outline the model of parental choice of the length of kindergarten attendance and derive the probability limit of the OLS and IV estimator. Additionally, I present the instrument (scarcity of kindergarten teachers) and assess its validity. In Chapter 3 I introduce the various data sources and discuss measurement of the most important variables, as well as the chosen set of control variables. In Chapter 4 I present descriptive statistics regarding the 1st stage, the 2nd stage and the reduced form relationship, as well as the spatial distribution of the instrument in Hungary. I investigate the relationship between the scarcity of kindergarten teachers and the quality of kindergarten service provision at municipality-level. I pay particular attention to whether the instrument is related to student achievement through channels other than time spent in kindergarten, like primary school quality. In Chapter 5 I present the OLS and IV estimates. I look at whether the results are sensitive to slight modifications of the instrument. I present estimates separately for parental education to explore any heterogeneous effects. In Chapter 6 I conclude and point out directions for further research.

[1] See for instance the following declaration (available only in Hungarian): http://www.kormany.hu./hu/nemzeti-eroforras-miniszterium/oktatasert-felelos-allamtitkarsag/hirek/harom-eves-kortol-kotelezove-tenne-az-ovodat-a-kormany.
According to the Hungarian Act No. LXXIX of 1993 on Public Education **24. §** (1), every child has to attend kindergarten for at least 4 hours per day in the year in which she becomes 5 years old.

Chapter 2: Theoretical Model and Identification

2.1. THE MODEL

Time spent in kindergarten can be thought of as an investment in a child's human capital. The effect of time spent in kindergarten on student achievement is a form of returns to schooling. In the language of program evaluation literature, variation of years spent in kindergarten can be thought of as variation in treatment. In this Chapter I present a theoretical model of the decision of time spent in kindergarten. Then I discuss how the measurement of the causal effect of time spent in kindergarten on student achievement fits into the standard literature of returns to education.

Measuring the causal effect of an additional year spent in school on an individual's lifetime earnings levels and patterns has been subject to intensive research among labor economists. In the standard problem every child starts school approximately at the same age and is obliged to stay in school for a minimum amount of time. Then, individuals choose the optimal number of years spent in formal education. Every individual decides for an additional year in school as long her internal rate of return[2] is higher than the interest rate. Whether this decision is made in advance or adjusted gradually might vary across models, but the margin is where to stop the schooling life cycle. The problem of how many years to spend in kindergarten is exactly the mirror image. Families decide about when their child should start kindergarten[3].

In this section I present a random coefficient model of the parental decision about time spent in kindergarten. As I mentioned above, this decision is analogous to the individual choice of time spent in school. The model I present is the application of Gary S. Becker's model by Willis (1987), developed by Card (1995, 1999, pp. 1810-1834) and refined by Heckman et al. (2006a, pp. 364-370). This is a static model that abstracts from sequential updating of information. After deriving the probability limit of the OLS and IV estimators, I discuss the Local Average Treatment Effect (LATE) interpretation of the IV estimator. Finally I present the particular instrument used in this paper, together with a discussion regarding its validity.

The following framework is exactly of that Card (1999, pp. 1810-1834) follows.

To keep the analysis as simple as possible, consider a one parent – one child family. The decision-maker is the parent, who has a utility function according to the following.

$$U(T_i, K_i) = B\left[T_i(K_i)\right] - C_i(K_i) \qquad (1)$$

[2] The internal rate of return equalizes the present value of two earnings streams net of direct costs of education associated with different schooling alternatives.

[3] Additionally, there is a second margin. Parents have to decide whether their child should stay in kindergarten at the age of 6, provided that she is allowed to start primary school. Both the model and the empirical part will focus on the first margin and keeps the second one fixed.

where $T_i(K_i)$ represents school test scores of child . This utility function depends positively on benefits from the child's test score, which is a function of time spent in kindergarten, K_i . It depends negatively on the costs, $C_i(K_i)$ stemming from the child's kindergarten attendance.[4]

The way how years spent in kindergarten transform into test scores varies across families, according to the following test score production function.

$$T_i(K_i) = b_{0i} + b_{1i}K_i - b_2 K_i^2 \tag{2}$$

where $b_{0i,} b_{1i}, b_2 > 0, \forall i$. Thus, I assume student achievement to be concave is time spent in kindergarten.

Costs are assumed to be convex in time spent in kindergarten and have the following form.

$$C_i(K_i) = c_0 + c_{1i}K_i + c_2 K_i^2 \tag{3}$$

where $c_0, c_{1i}, c_2 > 0 \forall i$.

Finally, I assume parents to be identical in the way they receive utility from the child's student achievement. In order to keep the argument simple, I assume function B to be the identity function. Substantive results carry forward as long as function B is a non-convex function and homogeneous across families.

Heterogeneity across families stem both from the benefit and the cost part of the utility function. Parental benefits are heterogeneous since the test score production function has both an individual-specific intercept and an individual-specific slope. These produce the random-coefficient/heterogeneous treatment effect model[5]. The slope of the cost function also varies across families. However, additional heterogeneity regarding the curvature of both the test score production function and the cost function is beyond the scope of this paper. I also leave the role and implications of individual-specific intercept in the cost function for further research.

Costs can be heterogeneous due to several reasons. First, due to scarcity or inferior quality of kindergarten services provided in the municipality that affects families to a different extent. Second, they might be heterogeneous due to various parental preferences. Disadvantaged minorities might fear of irreconcilable cultural differences experienced at their neighborhood *versus* experienced at kindergarten. They also might fear of the lack of cooperative attitude from kindergarten teachers, thus might have reservations about the care received by their child. Additionally, some parents might have difficulties with regular attendance in kindergarten. If parents are unemployed and do not have a regular daily routine, bringing the child to the kindergarten might cause difficulties. Unemployed parents are also generally the first being rejected in case of scarcity of local kindergarten services[6].

---

[4] In the original model Card first assumes a simple specification for the heterogeneity components in the marginal revenue function and then due to integration he gathers the implied model for log of earnings (Card, 1999, pp. 1811-1813).

[5] See for instance Björklund and Moffit (1987) for first discussion of random coefficient models.

[6] According to the Hungarian Act No. LXXIX of 1993 on Public Education **65.** § (2), every child should attend kindergarten in the neighbourhood in which she lives or her parents work. Every institution is obliged to admit children who are 5 years old and live in the neighbourhood. They are prohibited to reject multiply disadvantaged children if they are at least 3 years old and would like to be admitted.

There is ample empirical evidence for the positive effects of an additional year spent in kindergarten being the largest for children stemming from the least advantageous families (e.g. Blau and Currie (2006) provide a comprehensive survey). Disadvantaged children can be assumed to have lower intercept than those stemming from privileged families since, at least on average, they receive less positive stimuli at home. Consequently, they would have lower student achievement in the absence of kindergarten care. These children benefit presumably more from an additional year spent in kindergarten, exactly because their home environment is less supportive.

Thus, intuitively we can expect that the individual-specific intercept of the marginal return and the individual-specific cost parameter are negatively correlated. In less privileged families unemployed parents are overrepresented who are disfavored in case of kindergarten service scarcity or lack daily routine. In case of minorities they are most likely reluctant in sending their child to kindergarten because of distrust. Presumably they are exactly those parents, whose knowledge about early childhood development, the salient importance of positive cognitive and emotional stimuli is deficient. Thus, they are less likely to provide these stimuli at home. On the other hand, we can also expect intuitively that that the individual-specific slope of the marginal return and the individual-specific cost are positively correlated.

To understand convex costs and concave returns better, consider the following example. Suppose time spent in kindergarten is measured in years and the parent has to decide whether her child should start kindergarten at age 3 or at age 4 before she starts school at age of 6. Kindergarten at age 5 is compulsory. Suppose there are no breaks in kindergarten attendance. As I mentioned above, this situation is exactly the mirror image of the standard problem of how many years to spend in school at the end of the schooling life cycle. By sending the child to kindergarten at age of 3 (4) the parent implicitly decides for 3 (2) years of kindergarten attendance. Thus, in the first case the child spends 3 years in kindergarten, in the second 2.

Convex costs mean in general that the cost of the 3^{rd} year of attendance are higher than the costs of the 2^{nd} year of attendance, and costs of the 2^{nd} year are higher than costs of the 1^{st} year of attendance. In case of the mirror image problem of kindergarten attendance convex costs mean that the costs of the 3^{rd} year spent in kindergarten are the highest, which is practically the 1^{st} one between the child's age of 3 and 4. This happens e.g. if parents get more and more accustomed to bringing their child to kindergarten in their daily routine and feel less and less reluctant to bring their child to kindergarten as the compulsory year approaches.

Concave returns in general mean that the return of schooling declines with an additional year spent in school. In case of the mirror image kindergarten attendance concave returns mean that the returns of kindergarten attendance are the highest between age 5 and age 6, are lower between age 4 and age 5 and are the lowest between age 3 and age 4. According to Cunha et al. (2006), cognitive and non-cognitive skill formation is a cumulative process in which skills acquired at one stage in the development process enhance skill formation at later stages. Hence, while skills acquired at between age 3 and 4 contribute the most for further development, the returns might not be present immediately. Or, critical and sensitive periods in child development of skills emphasized by Heckman (2007) that are measured in school might be such that the return is the highest between age 5 and 6.

Although the intuition behind neither convex costs nor concave benefits is completely straightforward, at least one of these assumptions is necessary to have an interior optimum. The exact value of neither curvature alters the substantive results as long as they are identical across families.

To sum up, the parental utility function takes the following form by assuming that one test score translates into one unit of parental utility.

$$U(T_i, K_i) = b_{0i} + b_{1i}K_i - b_2 K_i^2 - (c_0 + c_{1i}K_i + c_2 K_i^2) \tag{4}$$

Every parent maximizes this utility function with respect to time spent in kindergarten An optimal choice satisfies the usual the first-order condition that equalizes marginal utility and marginal cost of kindergarten attendance, as in Equation (5). Equation (6) shows the optimal demand for kindergarten services (measured in time) for child *i*.

$$MR_i = b_{1i} - 2b_2 K_i = c_{1i} + 2c_2 K_i = MC_i \tag{5}$$

$$K_i^* = (b_{1i} - c_{1i}) / (2(b_2 + c_2)) \equiv (b_{1i} - c_{1i}) / k \tag{6}$$

Optimal demand for kindergarten services measured in time is linear in the individual-specific heterogeneity terms. It depends positively on the individual-specific slope of the marginal benefit and negatively on the individual-specific part of the marginal cost function. The higher the curvature of both the cost and benefit functions, the lower the optimal demand for kindergarten. Note that if both the marginal return and the marginal cost are independent of the level of time spent in kindergarten, there is no optimal kindergarten choice. It can be the case that either the benefit or the cost is linear in time spent in kindergarten, but it cannot be both.

2.2. Derivation of the probability limit of the OLS estimator

From equation (2) the average marginal effect of time spent in kindergarten on test score in the population is

$$E(b_{1i} - 2b_2 K_i) = b_1 - 2b_2 K \tag{7}$$

where b_1, K denote the mean individual-specific part of the marginal benefit and the mean time spent in kindergarten in the population, respectively. If every child's life path could be observed in case of different treatment intensities, then one could obtain the average effect in (7) of spending *k* years *versus k-1* years in kindergarten on test scores by averaging the differences between test scores obtained after *k* years *versus k-1* years calculated for every child. The identification problem arises due to the existence of several potential outcomes, but only one realized, thus for the researcher observable outcome per child. If time spent in kindergarten would be random in the population, then this average effect could be identified by a simple OLS. However, time spent in kindergarten is almost certain to be nonrandom. Analogously to what has been stressed in the standard problem of returns to education, the key endogeneity problem arises due to selection on unobservables. Are higher test scores observed for children having spent more years in kindergarten caused solely by their greater amount of pre-school care? Or does the parent decide systematically for different amount of kindergarten care not necessarily based on, but related to the child's benefit, or skills acquired from pre-school parental care? If parents make treatment choices related to heterogeneity in returns and if some components of the heterogeneity remain unobserved, then the question can be analyzed in a framework that Heckman et al. (2006b) label 'model with essential heterogeneity'.

Since OLS identifies average effects, it is useful to re-write Equation (2) the following way.

$$T_i(K_i) = b_{0i} + b_{1i}K_i - b_2K_i^2 = a_0 + b_1K_i - b_2K_i^2 + a_i + (b_{1i} - b_1)K_i \tag{8}$$

where $a_i = b_{0i} - a_0$, it has zero mean and $a_i + (b_{1i} - b_1)K_i$ is the residual. Equation (8) together with (6) together describes a two-equation system for time spent in kindergarten in terms of the underlying random variables a_i, b_{1i}, c_{1i}. Equation (9) shows the probability limit of the OLS estimator from regressing test score on time spent in kindergarten.

$$p\ lim\ b_{OLS} = \frac{Cov(T_i(K_i), K_i)}{Var(K_i)} = \frac{Cov(a_0 + b_1K_i - b_2K_i^2 + a_i + (b_{1i} - b_1)K_i, K_i)}{Var(K_i)} \tag{9}$$

Substituting the optimal time spent in kindergarten, represented by Equation (6) into (9), one can obtain the following expression for the probability limit of the OLS estimator.

$$p\ lim\ b_{OLS} = b_1 - 2b_2K^* + \mu_0 + \varphi_0K^* \tag{10}$$

where μ_0, φ_0 are obtained from the following linear projections and K^* denotes the population mean of optimal time spent in kindergarten.

$$\begin{aligned} a_i &= b_{0i} - a_0 = \mu_0(K_i^* - K^*) + u_i \\ b_{1i} - b_1 &= \varphi_0(K_i^* - K^*) + \upsilon_i \\ K_i^{*2} &= \vartheta_o K_i^* + \varsigma_i \end{aligned} \tag{11}$$

$E(u_i) = E(\upsilon_i) = E(\varsigma_i) = 0$ by definition of $\mu_0, \varphi_0, \vartheta_0$. These coefficients are practically OLS regression coefficients derived below (see Equations (12)).

$$\mu_0 = \frac{Cov(a_i, K_i^*)}{Var(K_i^*)} = k\frac{Cov(a_i, b_{1i} - c_{1i})}{Var(b_{1i}) + Var(c_{1i}) - 2Cov(b_{1i}, c_{1i})} = k\frac{Cov(a_i, b_{1i}) - Cov(a_i, c_{1i})}{Var(b_{1i}) + Var(c_{1i}) - 2Cov(b_{1i}, c_{1i})}$$

$$\varphi_0 = \frac{Cov(b_{1i}, K_i^*)}{Var(K_i^*)} = k\frac{Cov(b_{1i}, b_{1i} - c_{1i})}{Var(b_{1i}) + Var(c_{1i}) - 2Cov(b_{1i}, c_{1i})} = k\frac{Var(b_{1i}) - Cov(b_{1i}, c_{1i})}{Var(b_{1i}) + Var(c_{1i}) - 2Cov(b_{1i}, c_{1i})}$$

$$\vartheta_0 = \frac{Cov(K_i^{*2}, K^*_i)}{Var(K^*_i)} = \frac{E(K_i^{*3}) - E(K_i^{*2})E(K_i^*)}{E((K_i^* - K^*)^2)} = \frac{E(K^{*3}_i) - E(K_i^{*2})E(K_i^*)}{E(K_i^{*2}) - (E(K_i^*))^2} = 2K^* \tag{12}$$

where the following relationship between moments about the origin and moments about the mean is used: $E((K_i^* - K^*)^2) = E(K_i^{*2}) - (E(K_i^*))^2$ and $0 = E(K_i^{*3}) - 3E(K_i^*)E(K_i^{*2}) + 2(E(K_i^*)^3)$ with the assumption that $b_{1i,}c_{1i}$ have jointly symmetric distribution, thus $E((K_i^* - K^*)^3) = 0$.

Equation (10) is obtained using Equation (13), together with the assumption that $b_{1i,}c_{1i}$ have jointly symmetric distribution and substituting in φ_0 from (11).

$$
\begin{aligned}
&Cov((b_{1i}-b_1)K_i^*,K_i^*)=Cov((b_{1i}-b_1)K_i^*,K_i^*-K^*)=\\
&=E((b_{1i}-b_1)K_i^*(K_i^*-K^*))-E((b_{1i}-b_1)K_i^*)E(K_i^*-K^*)=\\
&=E((b_{1i}-b_1)K_i^*(K_i^*-K^*))=E((b_{1i}-b_1)(K_i^*-K^*+K)(K_i^*-K^*))=\\
&=E((b_{1i}-b_1)(K_i^*-K^*)(K_i^*-K^*))+E((b_{1i}-b_1)(K^*)(K_i^*-K^*))=\\
&=E(\varphi_0(K_i^*-K^*)(K_i^*-K^*)(K_i^*-K^*))+E(\varphi_0(K_i^*-K^*)(K^*)(K_i^*-K^*))=\\
&=\varphi_0 Var(K_i^*)K^*
\end{aligned}
\tag{13}
$$

The average marginal return of years spent in kindergarten on test scores in the population is thus $b_1 - 2b_2K^*$.

I refer to μ_0 as the bias stemming from 'sorting on the level'. Clearly, if there is no relationship between the individual-specific intercept of the marginal return of kindergarten attendance and optimal time spent in kindergarten, then this bias is not present. Previous discussion argued that disadvantaged children benefit more from additional unit of time spent in kindergarten, exactly because their home environment is less supportive. Consequently, $Cov(a_i, b_{1i}) < 0$. If marginal cost of kindergarten attendance is higher for less privileged families, then $Cov(a_i, c_{1i}) < 0$. The ratio of the two covariances, thus the two correlations is positive. Since

$$Cov(a_i,b_{1i})-Cov(a_i,c_{1i})=st(a_i)\left[Corr(a_i,b_{1i})sd(b_{1i})-Corr(a_1,c_{1i})sd(c_{1i})\right]$$

$$\mu_0 < 0 \text{ if } Corr(a_i,b_{1i})sd(b_{1i})-Corr(a_1,c_{1i})sd(c_{1i})<0 \tag{14}$$

Recall that b_{1i} measures the benefit effect of an additional unit of time spent in kindergarten on test score, thus on parental utility and c_{1i} measures the cost effect of an additional unit of time spent in kindergarten on parental utility (if both benefits and costs are linear in time spent in kindergarten).

Suppose that their variation is identical. Then the bias stemming from sorting on the level is downward if $Corr(a_i, c_{1i}) > Corr(a_i, b_{1i})$. Since I assume these correlations to be negative, sorting on the level produces a downward bias in the OLS estimates if ability brought from home is more related to benefits than to costs. If the variation in benefits is lower than the variation in the costs, then sorting on the level leads to a downward bias if ability brought from home is *even more* related to marginal benefits than to costs. If the variation in benefits is higher than in costs then the relationship between the two correlation coefficients is uncertain to produce a downward bias.

On the other hand,

$$\mu_0 > 0 \text{ if } Corr(a_i,b_{1i})sd(b_{1i})-Corr(a_1,c_{1i})sd(c_{1i})>0 \tag{15}$$

Suppose again that marginal costs and benefits are identically spread out. Then the bias stemming from sorting on the level is upward if $Corr(a_i, b_{1i}) > Corr(a_i, c_{1i})$. Since I assumed these correlations to be negative, sorting on the level produces an upward bias if ability brought from home is more related to costs than to benefits. If the variation in costs is lower than the variation in benefits, then sorting on the level leads to an upward bias if ability brought from home is *even more* related to marginal costs than to benefits. If the variation in costs is higher than in benefits then the relationship between the two correlation coefficients is ambiguous to produce an upward bias.

In the baseline problem of returns to education the sign of the bias stemming from sorting on the level is more predictable. In that problem it is a plausible assumption that the relationship between innate ability and heterogeneous part of the marginal return $Cov(a_i, b_{1i})$ is positive and $Cov(a_i, c_{1i})$ is similarly negative (Card (1999, pp. 1814)). This means that less able individuals who would earn lower wages keeping educational level fixed would presumably benefit less from an additional year of education at the end of schooling life cycle. Additionally, marginal costs of schooling can be assumed to be higher for less able persons.

To sum up, in case of returns to time spent in kindergarten the sign of the sorting on the level bias is ambiguous, while in the standard problem of schooling choice it is more predictable. If innate ability is sufficiently more (negatively) related to marginal benefit than to cost then OLS estimates are downward biased. This happens if skills acquired at before kindergarten are sufficiently strongly related to potential benefits from human capital investment in the kindergarten. According to the life cycle skill formation literature and Heckman (2007), gaps can be reduced if intervention takes place at early stages in the life cycle. Further research is needed to assess whether kindergarten care is a sufficiently early intervention to produce a large enough coariation $Cov(a_i, b_{1i})$ in absolute value. At the same time, this is a direct relationship, while $Cov(a_i, c_{1i})$ is rather a result of the relationship between family background and ability on the one hand and family background and marginal cost on the other. Indirect relationships may be weaker. Thus, I consider the downward bias case slightly more likely, if not ambiguous.

I refer to $\varphi_0 K$ as the bias stemming from 'sorting on the gain', the interdependence between the individual-specific slope and years spent in kindergarten.

$\varphi_0 > 0$ if $Var(b_{1i}) - Cov(b_{1i}, c_{1i}) > 0$, which is equivalent to

$$\frac{sd(b_{1i})}{sd(c_{1i})} > Corr(b_{1i}, c_{1i}) \tag{16}$$

If the relationship between the individual-specific part of the marginal cost and marginal benefits is negative, then the bias stemming from sorting on the gain is unambiguously upward. This happens if those who have lower marginal cost benefit more from an additional year spent in kindergarten.

However, before I argued that children from a less privileged family gain more from additional year spent in kindergarten, but occur higher costs. Thus, it can be assumed that $Corr(b_{1i}, c_{1i}) > 0$ .Thus, the marginal benefit can be more or less spread out than marginal cost in order to sorting on the gain leading to an upward bias.

On the contrary,

$\varphi_0 < 0$ if $Var(b_{1i}) - Cov(b_{1i}, c_{1i}) < 0$, which is equivalent to

$$\frac{sd(b_{1i})}{sd(c_{1i})} < Corr(b_{1i}, c_{1i}) \tag{17}$$

Taking into account that a correlation coefficient is always between zero and one, a necessary but not sufficient condition for sorting on the gain producing a downward bias is that the variation in the marginal cost is larger than the variation in the marginal benefit. As I mentioned above, if the variation in benefits is lower than the variation in the costs, then sorting on the level leads to a downward bias if ability brought from home is *even more* related to marginal benefits than to costs. Thus, sorting on the level and the gain are biased towards the same direction, downwards, under the aforementioned assumptions.

In the standard problem of returns to education the direction of the bias is more predictable. There it is argued that if individuals know, or are able to predict at least partially their own returns at the time they make their schooling decisions, those with higher return to schooling will self-select themselves into more schooling. This leads to an upward bias in the estimated coefficient from a cross-sectional regression of earnings on schooling, even in the absence of variation in unobserved raw ability and produces the 'correlated random coefficient model'. In the standard problem it is assumed that the relationship between marginal cost and marginal benefit is negative, thus sorting on the gain produces an unambiguous upward bias (see e.g. Card (1999, pp. 1813)).

To assess the magnitude of the potential bias stemming from sorting on the gain, consider the fraction of variance of years spent in kindergarten attributable to variation in the individual-specific slope (as opposed to tastes and costs). This is shown in Equation (18).

$$\frac{Var(b_{1i}) - Cov(b_{1i}, c_{1i})}{Var(b_{1i}) + Var(c_{1i}) - 2Cov(b_{1i}, c_{1i})} \tag{18}$$

This expression is smaller than one if

$$Cov(c_{1i}, c_{1i} - b_{1i}) > 0 \leftrightarrow Cov(c_{1i}, b_{1i} - c_{1i}) < 0 \tag{19}$$

Thus, the fraction of variance of years spent in kindergarten attributable to variation in the individual-specific slope is smaller than one if those with higher marginal cost obtain systematically less kindergarten education. This is a very safe assumption. Thus, the larger the aforementioned fraction, the larger the endogeneity bias. This result is the same as obtained in the standard problem of returns to schooling (Card (1999, pp. 1815)).

2.3. DERIVATION OF THE PROBABILITY LIMIT OF THE IV ESTIMATOR

IV estimation might help to overcome endogeneity bias. Every valid instrument should, irrespective of homogeneous or heterogeneous treatment effects, represent an exogenous variation behind the choice regarding time spent in kindergarten and satisfy several conditions. First, the instrument should be uncorrelated with disturbances obtained both from the relationship of interest and the years spent in kindergarten equation. Second, the instrument should significantly affect treatment. This means that its regression coefficient should be significantly different from zero in the years spent in kindergarten equation. In the framework of two-staged least squares (2SLS) estimation the time spent in kindergarten equation is the 1st stage. The predicted value of the 1st stage can be substituted into the 2nd stage, the relationship of interest.

Taking into account that the instrument is explicitly excluded from the 2nd stage equation, the intuition behind IV estimation is straightforward. The effect of the instrument operates only through the endogenous variable of interest, but has no direct effect on the outcome of interest. The intuition behind the exclusion restriction is that only one particular part of the variation of years spent in kindergarten is used, exactly that part that is associated with the instrument. Subsequently, the only reason why test score varies with the instrument is that years spent in kindergarten varies with the instrument.

If returns are homogeneous then all valid instruments define the same population parameter. On the contrary, if returns are heterogeneous then IV estimates might be different from the population effect. If treatment adoption is made without knowledge about the idiosyncratic gain from treatment, then IV identifies the mean population response parameter. However, if treatment adoption is made with partial or full knowledge about treatment, then IV will not identify the mean population response parameter (Heckman et al., 2006b).

Angrist and Krueger mention two possible sources of instruments in case of the standard schooling choice (2001, pp. 73). In the first case exogenous variation stems from differences in costs that vary independently of ability, tastes and earnings potential. In the second case exogenous variation comes from institutional constraints. For instance, college proximity can be thought of as a source of differences in costs across individuals (Card, 1995b), while compulsory school attendance laws represent institutional constraints (Angrist and Krueger, 1991).

In the present analysis, the instrument stems from differences in costs that (hopefully) vary independently of ability, tastes and achievement potential. I assume scarcity of kindergarten service to generate such exogenous supply-side constraint.[7] Families experiencing scarcity of kindergarten service in the municipality they live might be forced to commute to other municipalities or use resources for lobbying for their child's admission.

Scarcity of kindergarten service could mean, among others, shortage in relative number of kindergarten places or kindergarten teachers, or in relative amount on kindergarten public expenditures. I will use the number of kindergarten teachers in the municipality over the 3-5 year aged population in the municipality at the child's age of 4 as a supply-side constraint. I use the number of kindergarten teachers instead of kindergarten places because presumably there is less exogeneous variation behind the creation and termination of kindergarten places than behind the hiring or layoff of kindergarten teachers that causes scarcity.

In particular, I will use this measure in a binary form. I will specify the cutoff-value later that determines whether the number of teachers happened to be especially low at the family's municipality at the child's age of 4, meaning serious scarcity of kindergarten teachers. It is important to note that I will use information from both the extensive and the intensive margin of kindergarten service provision. Thus, scarcity of kindergarten teachers will apply for municipalities with the lowest relative number of teachers and for those with no kindergarten service provision at all.

[7] Note that the distinction between the two sources of the instruments is not straightforward. Thus, scarcity of kindergarten teachers is not an instrument that stems unambiguously from the cost side. If every child were obliged to attend kindergarten in the neighborhood he/she lives, then scarcity of kindergarten teachers could stem from institutional constraints, as well.

The presumed mechanism is the following. First, scarcity of kindergarten teachers affects the number of time spent in kindergarten. I assume that those who lived in municipalities at their age of 4 where there happened to be a serious shortage of kindergarten teachers have spent systematically less time in kindergarten. Second, I assume that scarcity of kindergarten teachers has no effect on test score other than that operates through kindergarten attendance. If these requirements are fulfilled, then this instrument is a valid one.

To formalize the aforementioned argument, suppose that the individual-specific marginal cost component is linearly related to a Z_i variable: $c_{1i} = \pi_i Z_i + \varepsilon_i$. Hence, Equation (6) can be re-written as the following.

$$K_i = \pi_0 + \pi Z_i + \omega_i \tag{20}$$

where $\omega_i = (b_{1i} - b_1 - \varepsilon_i) / k$. Using Z_i as an instrument, IV estimation will provide a consistent estimate for the mean return to kindergarten care in the population if the followings assumptions are valid (Card, 1999, pp. 1817, based on Wooldridge, 1997):

$$E(\varepsilon_i \mid Z_i) = E(a_i \mid Z_i) = E((b_{1i} - b_1) \mid Z_i) = 0 \tag{21}$$

$$E((b_{1i} - b_1)^2 \mid Z_i) = \sigma_b{}^2 \tag{22}$$

$$E(\omega_i \mid b_{1i}, Z_i) \text{ is linear in } b_{1i}. \tag{23}$$

The first assumption (Equation (21)) is that all individual-specific terms (the intercept, the slope and the residual of the marginal cost after partialling out the effect of the instrument) are mean-independent of the instrument. Thus one seeks for an instrument that is mean-independent of innate ability components, the return and any determinant of the marginal cost apart from the instrument. The second assumption (Equation (22)) says that the second moment of the individual-specific return, thus its spread is conditionally independent of the instrument.

Now I am deriving the probability limit of the IV estimator. For simplicity and in favor of the specification I will present later, I assume that benefits are linear in time spent in kindergarten.

$$p\ lim\ b_{IV} = \frac{Cov(T_i(K_i), Z_i)}{Cov(K_i, Z_i)} = \frac{Cov(a_0 + b_1 K_i + a_i + (b_{1i} - b_1)K_i, Z_i)}{Cov(K_i, Z_i)} \tag{24}$$

Substituting the optimal time spent in kindergarten, represented by Equation (6) into (25), one can obtain the following expression for the probability limit of the IV estimator.

$$p\ lim\ b_{IV} = b_1 + \mu_0{}' + \varphi_0{}' K^* \tag{25}$$

where μ_0, φ_0 are obtained from the following projections and K^* denotes the population mean of optimal time spent in kindergarten.

$$\begin{aligned} a_i = b_{0i} - a_0 &= \mu_0(K_i^* - K^*) + u_i \\ b_{1i} - b_1 &= \varphi_0(K_i^* - K^*) + \upsilon_i \end{aligned} \tag{26}$$

$E(u_i) = E(\upsilon_i)$ by construction and

$$\mu_0' = \frac{Cov(a_i, Z_i)}{Cov(K_i^*, Z_i)} = c_2 \frac{Cov(a_i, Z_i)}{Cov(b_{1i}, Z_i) - Cov(c_{1i}, Z_i)} \quad (27)$$

$$\varphi_0' = \frac{Cov(b_{1i}, Z_i)}{Cov(K_i^*, Z_i)} = c_2 \frac{Cov(b_{1i}, Z_i)}{Cov(b_{1i}, Z_i) - Cov(c_{1i}, Z_i)} \quad (28)$$

Equation (25) is obtained using Equation (29), together with the assumption that $b_{1i,}c_{1i}$ have jointly symmetric distribution and substituting in φ_0' from (28).

$$\begin{aligned}
& Cov((b_{1i} - b_1)K_i^*, Z_i) = Cov((b_{1i} - b_1)K_i^*, Z_i - Z) = \\
& = E((b_{1i} - b_1)K_i^*(Z_i - Z)) - E((b_{1i} - b_1)K_i^*)E(Z_i - Z) = \\
& = E((b_{1i} - b_1)K_i^*(Z_i - Z)) = E((b_{1i} - b_1)(K_i^* - K^* + K^*)(Z_i - Z)) = \\
& = E((b_{1i} - b_1)(K_i^* - K^*)(Z_i - Z)) + E((b_{1i} - b_1)(K^*)(Z_i - Z)) = \\
& = E(\varphi_0'(K_i - K^*)(K_i^* - K^*)(Z_i - Z)) + E(\varphi_0'(K_i^* - K^*)(K^*)(Z_i - Z)) = \\
& = \varphi_0' Cov(K_i^* Z_i)K^*
\end{aligned} \quad (29)$$

The average marginal return of years spent in kindergarten on test scores in the population is b_1. Before I investigate the terms μ_0', φ_0' in more detail, it is useful to make the distinction between the two channels through which the instrument, scarcity of kindergarten teachers can be related to the individual-specific parameters.

First, it can happen that scarcity of kindergarten teachers is not randomly distributed across space, but is concentrated in municipalities in which less privileged families are overrepresented. I argued before that children from less privileged families can be assumed to have lower a_i, but higher b_{1i}, c_{1i}. Consequently I assume that $Cov(a_i, Z_i) < 0, Cov(b_{1i}, Z_i) > 0, Cov(c_{1i}, Z_i) > 0$. Denote this channel of relationship between the instrument and the individual-specific parameters 'endogeneity channel'.

Second, the instrument itself can have an effect on the individual-specific parameters through the mechanism. Scarcity of kindergarten teachers increases kindergarten attendance costs to every family to some extent. Families with different family background have different costs. If kindergarten scarcity increases the cost of additional year of kindergarten attendance for disadvantaged families more than for privileged families then the covariance between individual-specific cost and scarcity of kindergarten teachers is positive. This happens if in case of scarcity of teachers disadvantaged minority children or children with unemployed parents are more likely to be rejected.

Scarcity of kindergarten teachers might also lead to larger kindergarten class sizes which reduces benefits from kindergarten attendance to every family to some extent. Consider two children, the one receives a considerable amount of positive stimuli at home, the second does not. Both of them have positive benefits from kindergarten attendance and in case of crowdedness, both of them get lower benefits. If the decrease in the benefit due to crowdedness is lower for the second child, then the covariance between individual-specific benefit and scarcity of kindergarten teachers is negative. This happens if the disadvantaged child needs more care to outweigh the negative effect of insufficient positive stimuli received at home.

There is no intuition behind why scarcity of kindergarten teachers would have an effect on innate ability acquired at home. I will refer to this second channel of the effect of the instrument on individual-specific parameters as ‘part of the mechanism’.

Now I can turn to the discussion of the two sources of bias in the IV estimates.

μ_0' is referred to as the bias stemming from ‘sorting on the level’. Based on the previous argumentation I assume that $Cov(a_i, Z_i) < 0$. Hence, $\mu_0' > 0$ if

$$Cov(b_{1i}, Z_i) - Cov(c_{1i}, Z_i) < 0 \Leftrightarrow sd(b_{1i})Corr(b_{1i}, Z_i) - sd(c_{1i})Corr(c_{1i}, Z_i) < 0 \quad (30)$$

The requirement in (30) for sorting on the level producing an upward bias can be justified if $Cov(b_{1i}, Z_i)$ is small enough and $Cov(c_{1i}, Z_i)$ large enough. As assessed above, $Cov(b_{1i}, Z_i)$ is likely to be positive through the endogeneity channel but negative through the mechanism. On the contrary, $Cov(c_{1i}, Z_i)$ is likely to be positive through both channels. Hence, it is very likely that sorting on the gain produces an upward bias, unless the covariance between the instrument and the marginal benefit is too large.

I refer to φ_0' as the bias stemming from ‘sorting on the gain’. $\varphi_0' > 0$ if

$$Cov(b_{1i}, Z_i) < 0 \text{ and } Cov(b_{1i}, Z_i) - Cov(c_{1i}, Z_i) < 0 \quad (31)$$

The first part, the relationship between the individual-specific marginal benefit component and scarcity of teachers is negative if the endogeneity channel is the inferior compared to what happens part of the mechanism. Thus, the relationship is negative if scarcity of teachers affects benefits through kindergarten quality to a greater extent than scarcity of teachers is related to benefits through family background and spatial distribution of families. If this relationship is negative then $Cov(b_{1i}, Z_i) - Cov(c_{1i}, Z_i) < 0$ follows from previous argumentation.

The bias stemming from sorting on the gain is downward if

$$Cov(b_{1i}, Z_i) > 0 \text{ and } Cov(b_{1i}, Z_i) - Cov(c_{1i}, Z_i) > 0 \quad (32)$$

However, previously I showed that the difference between the two covariances is likely to be positive. Therefore I consider sorting on the gain producing a downward bias to be very unlikely.

Consequently, putting the bias stemming from sorting on the level and sorting on the gain together, the IV estimator is very likely to estimate an effect that is higher than the average population effect.

Imbens and Angrist emphasize that for defining any average treatment effect in case of heterogeneous returns it is insufficient for the instrument to be independent of individual potential outcomes and potential treatment intensities, but related to the treatment indicator (1994, pp. 469). The identification problem arises through the possibility that although the causal effect might be strictly positive for every individual, the size of switchers into and out of treatment can be such that the causal effect for switchers into non-treatment might be outweighed by the effect for switchers into the opposite direction. One possibility to avoid this situation is to prevent two-way flows into and out of treatment by imposing a nonparametric restriction on how the instrument affects treatment. Through the monotonicity assumption it is required that although the instrument may not affect some individuals, every affected individual responds to changes in the instrument into the same direction in a monotone way[8]. In other words, the assumption requires that in case *every* individual is more likely to be treated at a particular value *a* of the instrument than another *b* value, then *any* individual who would be treated conditional on *b* is treated conditional on *a*.

If the aforementioned independence and monotonicity conditions are satisfied, then LATE can be estimated. In that case LATE is by definition the average causal effect of treatment for individuals whose treatment status if affected by the change in the instrument. If the instrument represents a supply-side constraint as the existence of nearby colleges, then LATE identifies the treatment effect for whom this constraint is most binding. These individuals were induced to attend college if they would have been lived in geographical proximity of it. LATE is the discrete approximation of the marginal treatment effect introduced by Björklund and Moffitt (1987), that is a willingness to pay measure for an infinitesimal amount of additional schooling.[9]

In the context of the present analysis, the monotonicity assumption is the following. Families living in a municipality experiencing abundance of kindergarten teachers are required to decide for at least as much time spent in kindergarten as they would have decided had they been lived in a municipality with scarcity. Clearly, this assumption is violated only if some families have very unusual preferences for kindergarten. Therefore it is a much more innocuous assumption than the validity of this instrument which I discuss at the end of this chapter.

Angrist and Imbens (1995) show how the method of 2SLS can be used to estimate average causal effects in the presence of multiple treatment intensity and non-exogenous assignment to treatment. They show that under the aforementioned independence and monotonicity assumptions and provided that the Stable Unit Treatment Values Assumption (SUTVA) is satisfied[10], the average causal response is a weighted average of the individual-specific slopes along a response function. This is in line with Heckman et al. (2006a, chapter 8.) who argue that the IV estimate can be obtained by the weighted average of MTE[11]. Weights are proportional to the number of affected individuals who may or may not be

[8] Other possibilities include assuming constant treatment effect for all individuals or assuming the existence of a particular value of the instrument, so that the probability of treatment conditional on that value is zero (Imbens and Angrist, 1994, pp. 469). Exactly this monotonicity assumption is heavily criticized by Heckman et al. (2006a, pp. 377): they argue that it is a too strong assumption that rules out heterogeneity in the response of schooling choices to the instruments while it allows for heterogeneity of responses to schooling. They argue that this assumption cannot be justified in many dynamic discrete choice models of schooling. However, it is always satisfied for latent index models for endogenous treatment, as it has been pointed out by Imbens and Angrist (1994, pp. 469).

[9] As Heckman et al. (2006b, pp. 392) point out, the value of LATE is dependent on the particular instrument that has been used for estimation. Consequently, by having more instruments that are both valid and strong does not in general lead to an improvement of the targeted parameter, the average response, compared to OLS. Clearly, this problem arises only with heterogeneous responses.

[10] The Stable Unit Treatment Assumption (SUTVA) rules out any general equilibrium effect by assuming no interference between the individuals. In particular it assumes that potential (or counterfactual) outcomes of a particular individual are independent of the treatment of any other individual, and this applies to every individual.

[11] The authors define several treatment effects as the policy relevant treatment effect (PRTE) which is the difference between aggregate per capita outcomes under the alternative and the baseline policies, the treatment on the untreated (TUT) or treatment on the treated (TT). They point out that if the instrument is exactly the policy one wishes to evaluate then the policy relevant treatment effect and the IV estimand coincide (Heckman et al., 2006, pp. 380). Additionally, they show that the average treatment effect (ATE), the TT, the TUT, the PRTE and the IV and OLS estimates are all weighted averages of the MTE, where the weights integrate to 1, but can be negative. They also show that the weights are different

representative of the entire population. This means that the weight attached to an individual is proportional to the magnitude of the effect of the instrument on him/her.

As Angrist and Imbens emphasize (1995, pp. 435), it is impossible to identify the affected individuals since their counterfactual treatment status cannot be observed. Subsequently, while LATE satisfies internal validity, it might not meet the criteria of external validity. However, in case the monotonicity assumption is satisfied, the causal response weighting function can help to gather some imagination about the size and attributes of groups that contribute to the average causal effect. This can be achieved by estimating the cumulative distribution function of the endogenous variable of interest with the instrument switched on an off (Angrist and Imbens, 1995, pp. 438).

Returning to the model outlined at the beginning of the chapter, consider a population that can be divided into G numbers of discrete subgroups. In each group g individuals have identical latent ability and cost terms ($a_g, b_{1g}, \varepsilon_g$). Suppose that the relative number of teachers decreases due to an intervention that leads to a change $\triangle K_g$ in the average years spent in kindergarten in any group g. Suppose furthermore, that children in the treatment and comparison group with identical latent ability and cost terms $a_g, b_{1g}, \varepsilon_g$ would spend the same years in kindergarten in the absence of the intervention and the joint distribution of abilities and costs are the same in the two groups. If B_g is the marginal return of years spent in kindergarten on primary school outcomes for any individual in group g in the absence of the intervention, then

$$p\ lim\ b_{IV} = \frac{E(B_g \triangle K_g)}{E(K_g)}$$

Sufficient conditions for IV estimation to identify the average marginal return in the population are for instance, identical marginal returns to education for each group g; or homogeneous additive treatment effect of the intervention, $\triangle K \forall g$.

Finally I am reviewing the potential threats to the validity of my chosen instrument and review the most important questions that the results of this chapter imply.

First, the chosen instrument might not be independent of tastes. Families are not randomly distributed across space, but, as a response to the local (government) service provision, they sort themselves into municipalities according to Tiebout-sorting. Suppose that a representative household's location choice affects not only its kindergarten quality consumption, but simultaneously its consumption of other local goods (for instance, public safety, job market opportunities, air quality, housing opportunities) as it has been assumed by Bayer (2000). Then the number of kindergarten teachers at the chosen municipality is also the result of both underlying differences in preferences for kindergarten services and other factors that shape the household's residential decision. However, mobility within Hungary is presumably too low in order to Tiebout-sorting causing severe estimation problems.

from the IV estimand and for the treatment parameters, thus in general, IV does not estimate the treatment parameters (Heckman et al., 2006a, Table 9).

However, even if mobility is especially low, local governments might adjust the supply of kindergarten services as a response to parental demand. In that sense the supply of these services, thus also the number of kindergarten teachers do not represent an exogenous variation behind kindergarten schooling choices. If a greater supply of kindergarten services is a result of stronger unobserved demand, higher kindergarten attendance can be falsely attributed to a better access to it. Stronger unobserved demand might be related to family background. Thus it is important to assess whether the spatial distribution of the instrument is related to the spatial distribution of family background.

Suppose for the moment, that scarcity of kindergarten services is exogenous for families and their demand for kindergarten services. If municipalities experiencing a serious shortage from kindergarten teachers do also provide inferior schooling quality in primary school, then this could lead to lower test scores independently of the relative number of kindergarten teachers. This is analogous to what has been stressed by Angrist and Krueger by (1999, pp. 1301). Random assignment alone does not provide an unambiguously valid instrument[12]. In this case, even if exogeneity in the reduced form equation can be assumed, it would be not true that the only channel through the instrument affects test scores is through the number of years spent in kindergarten. This might be true, even if there is substantial controversy in the educational production function literature over whether school resources matter for student achievement (see e.g. Todd and Wolpin (2007)). If it is true, then the IV-estimates are upward biased due to the negative correlation between scarcity of kindergarten teachers and primary school quality on the one hand and the negative correlation between scarcity and time spent in kindergarten on the other.

In this Chapter I derived the probability limit of the OLS and the IV estimator. I showed that IV estimates are likely to be upward biased if the instrument is not randomly distributed in space and is related to student achievement through primary school quality. The direction of the bias in the OLS estimates is more ambiguous, but it is slightly more likely to be downward biased.

As a natural next step, after data description and measurement issues I continue with descriptive statistics. I will show statistics regarding the 2nd stage relationship between time spent in kindergarten and student achievement, the 1st stage relationship between scarcity of kindergarten teachers and time spent in kindergarten and the reduced form relationship between scarcity of kindergarten teachers and average municipality-level test scores. I also present the spatial distribution of the instrument. I investigate how the instrument is related to measures that proxy for inferior kindergarten service quality. This is important since it is part of the mechanism and could decrease benefits of kindergarten attendance (as I mentioned increased class size as an example). Last, but not at least I will examine whether the instrument is correlated with other factors than kindergarten attendance that affect student achievement. Besides non-random distribution this is the other major concern regarding the validity of the instrument.

[12] Angrist and Krueger (1999, pp. 1301) mention the possibility that despite the number that determined whether one needed to serve in the military was chosen by lottery in their famous example of measuring the effect of veteran status on lifetime earnings, individuals with low draft-lottery numbers could continue their studies with a systematically higher probability in order to extend a draft deferment. If this is the case, then this induces a non-negligible relationship between their number (the instrument) and their wages (the outcome of interest), thus causing bias for the estimate of veteran status on lifetime earnings. The key here is that draft-eligibility had other consequences on lifetime wages rather than only influencing the probability of being a veteran.

Chapter 3: Data Description and Measurement

In this Chapter first I describe the various sources of the data used for empirical investigation. Then I describe the measurement of the most important variables and discuss some shortcomings of the data. Finally, I argue for the necessity of the chosen set of control variables and describe them.

In order to estimate the effect of time spent in kindergarten on student achievement, one needs data on kindergarten attendance and educational achievement. Additionally, one needs to observe important background characteristics that are related to both kindergarten attendance and student achievement. In order to construct the instrument, one needs to observe both the number of kindergarten teachers and the number of children with different ages at municipality-level for several years in Hungary. To assess whether there is a relationship between scarcity of kindergarten service provision and kindergarten service quality, data on municipality-level kindergarten service characteristics is needed. To assess whether the instrument is related to student achievement through other channels other than kindergarten attendance (especially primary school quality), data on municipality-level primary school service characteristics is needed.

I use data from 2008 for 6th graders on student achievement and kindergarten attendance, along with background characteristics from the Hungarian National Assessment of Basic Competences (HNABC)[13]. In general HNABC has been measuring the literacy and mathematics skills of 6th, 8th and 10th graders, and various basic competences such as counting, thinking, systematizing, and writing of 4th graders with a final and uniform concept from 2006. It is important to note that HNABC does not measure the students' knowledge regarding the compulsory curriculum. Rather it measures how the students are able to apply their acquired skills in realistic situations and to what extent they possess the necessary competences for further development.

I chose 6th graders in 2008 for various reasons. First, this cohort can be fully comprehensively observed in 2008. Data from 2006 and 2007 cannot be used due to technical reasons. In 2006 and 2007 the number of response categories to the question about how many years the child spent in kindergarten is insufficient. The response categories that have been offered were the following: none, less than 1 year, 1 year, between 1 and 2 years, 2 years, 3 or more years spent in kindergarten. Taking the distribution of years spent in kindergarten from 2008 for 8th graders into consideration, 48% of the children spent 3 and 40% 4 years in school[14]. By using the data from 2006 and 2007, one loses an important margin between years 3 and 4. In an empirical investigation that measures the effect of additional year spent in school, thus uses a linear specification losing this margin would be problematic[15]. Dropping year 2006 was attached by dropping the 4th graders cohort, since data for 4th graders was accessible only for 2006. I dropped year

[13] The goal of HNABC is multiple. First, it measures the achievement of the entire Hungarian educational system and provides national data for the governmental educational policies. Second, measures the achievement of various schools and provide data for local governmental educational policies. Finally, it serves as a tool for enhancing the practice of own measurement of school effectiveness for the individual schools.

[14] These figures were obtained by exclusion of particular children that will be specified below.

[15] It would not be problematic in a non-linear specification in which years spent in school enters in a form of several binary variables.

2009 because the goal of this paper is not to assess whether the relationship between test score and kindergarten attendance is stable in time (across years), but to measure the relationship at least for one cross-section reliably. Moreover, having only two data points in a time series dimension is insufficient to assess anything about stability in time. Hence, stability issues are left for further research. I dropped 10th graders in order to be able to concentrate solely on one educational type, primary school[16]. Finally, I dropped the 8th graders in 2008 to focus solely on one cross-section and to leave the stability of the relationship of interest across grades for further research.

Similar to other years and cohorts, 6th graders in 2008 had to complete a test sheet in mathematics and reading. The completion of the test sheets was mandatory for all children and the result of almost all students was centrally processed[17]. Exceptions were made to students with special needs, e.g. corporeally/sensually/mentally disabled or autistic students. Students who suffered from physic developmental deficits (e.g. suffered from behavioural problems or dyslexia/dysgraphia/dyscalculia) were required to complete the test sheet, but their result were not taken into account in calculating the school's achievement. Children who suffered from some kind of temporary injury that made them physically unable to do this were not required to complete the test sheet, and those who missed the class on the particular testing day.

I obtained data on background variables and kindergarten attendance from the student background survey from HNABC. The data set from this survey contains detailed information about the students' demographical and family background, but completion of the survey was non-compulsory.

The instrument is the number of kindergarten teachers over the kindergarten aged population in the municipality where the individual lived at her age of 4. I obtained data on the nominator of the instrument from KIR-Stat. KIR-Stat provides the most comprehensive data about the Hungarian educational system. Every educational institution every year is required to fill out a data form according to the 229/2006. (XI. 20.) government enactment. I aggregated institution-level data to municipality-level and then I matched to each observation the number of kindergarten teachers in the municipality she lived at her age of 4. I also obtained the number of kindergarten places for a robustness check.

I received the data for the denominator (number of kindergarten-aged population in the municipality) from Gábor Kertesi who asked the Hungarian Statistical Authority to compile a panel dataset of the population with different ages for each Hungarian municipality. I denote this data set DEM. I also use data for primary school-aged population for calculating the relative number of primary school teachers in the municipality.

Except for public expenditures, I obtained data about educational data for kindergarten and primary school from KIR-Stat. Table1 in the Appendix shows the variables I use to assess the quality of

[16] In the Hungarian educational system there are some institutions that provide a 6-year-long or an 8-year-long secondary education. Students can choose the 8-year-long track after the 4th grade and the 6-year-long track after the 6th grade. After the 8th grade at age 14, every child has to continue studies in one type of the secondary schools until age of 18. In the data set of 6th graders, 4.12% of the children without special education needs attended secondary school.

[17] There is some variation in the group of students whose result has been centrally processed across the years. For instance, in 2007 the 6th graders and in 2008 the 4th graders were required only to complete the test sheets in 200, nationally representative schools fully comprehensively. Or before 2008, 20 or 30 students were selected for testing from all plants of each institution and educational type for the 10th graders.

municipality-level kindergarten service and primary school service. I use qualification level of the teachers and the number of classes with different size. I use data for kindergartens from 1999 when the majority of the 6th graders in 2008 was 4-year old. I use data for primary schools for 2001 when the majority of the 6th graders in 2008 was 6-year old and started primary school.

Data about municipality-level public educational expenditure stems from the Local Governmental Treasury data base (LGT). Table1 in the Appendix shows that I use wage and non-wage expenditures per child, per teacher and per class. The content of the aforementioned data base changed substantially and regularly during the previous decades, and in 2001 data solely for primary school expenditures was not available. Therefore, I use expenditure for general education throughout the analysis. Expenditure for general education incorporates primary and general secondary school expenditures, but excludes vocational expenditures.

Time spent in kindergarten is measured in years with values of 0, 0.5, 1, 1.5, 2, 3 and 4. Scores in mathematics and reading are standardized test scores with zero mean and standard deviation of one. I define scarcity of kindergarten teachers from HNABC 2008 (and not from municipality-level data). I define for each individual a dummy variable of scarcity that equals 1 if the observation lived in a municipality at her age of 4 where the teacher over kindergarten-aged population ratio happened to be in the lowest sextile and 0 otherwise. Thus, scarcity of kindergarten teachers affects approximately 16% of students in the HNABC 2008 data base. For children who lived in municipalities where there was no kindergarten service provision at all, the value of this variable is automatically 1. I chose this value in order to define scarcity for a relatively low fraction of students. Otherwise there are chances that the instrument is binding only for very few individuals. Nevertheless, this cutoff-value is clearly arbitrary. I leave it for further research how sensitive the results are with respect to changing this cutoff-value. I provide some preliminary results as motivation for further research at the end of the thesis.

There are two main problems with the data that are likely to introduce bias in the estimates.

First, I cannot use information about all 6th graders in 2008. I select children who have no special educational needs, have valid test score either in reading or in mathematics and their kindergarten attendance and highest parental educational attainment can be observed. This means that 79802 observations remain from 107654 (74.1%). I standardize the test score to be a zero mean, one standard deviation variable for these 79802 observations. Table2 in the Appendix shows the details of the sample selection. The majority of excluded students does not have valid test score or did not complete the student background survey.

The excluded observations are very likely to be non-random. The missing response analysis of HNABC 2006 and 2007 (KOSTB102, KOSTB104[18]) reveals that among students who do not have valid test scores, those with lower parental educational attainment are overrepresented. Additionally, students with valid test scores obtained systematically higher grade from mathematics the year before. On an institution-level there is a significantly positive relationship between the average test scores and the

[18] These analyses were written in the Hungarian Academy of Sciences, Institute of Economics, Research Unit: Economics of Education as a research report . They can be downloaded from http://www.econ.core.hu/kutatas/edu/produktumok/kostb.html .

fraction of students who completed the background survey. There is also a significantly positive relationship between the average mathematics grades and the likelihood of completion the background survey. On an individual-level, one obtains a similarly positive relationship between average test scores and non-response behavior. Hence, among excluded students those with less privileged family background and worse student achievement are very likely to be overrepresented also in 2008. If the population effect is higher for less privileged students then the aforementioned selection likely causes a downward bias in my estimates.

Second, there is measurement error in the instrument, because I cannot observe the municipality the child lived at age of 4, only her residence in 2008. Moreover, I cannot observe whether the child indeed attended kindergarten in the municipality she lived at her age of 4. There were also municipalities that did not provide any kindergarten service; therefore, families have to commute for kindergarten attendance. It would lead to more precise measurement if I could calculate the scarcity of kindergarten teachers for groups of municipalities, based on individual commuting behavior. Measurement error leads in general to downward biased coefficients[19].

Regarding the control variables, in general one needs to keep those factors fixed that affect both the variable of interest and the outcome of interest. The exclusion of these variables would lead to omitted variable bias. Furthermore, inclusion of any control variable should not raise simultaneity concerns. For instance, one should not control for the number of books the child has on its own, since there is very likely a two-way relationship between number of books and literacy scores. On the one hand, children who have serious problems with reading and understanding, and have low literacy test scores are less likely to get books as presents. On the other hand, own books might further enhance literacy skills for those who have such skills above a certain level. Similarly, the mother's employment history should not be taken into account since it could have been affected by her child's kindergarten attendance.

In the case of IV estimation, one needs to pay special attention to the estimation of the 1st stage. As the instrument is defined for every child's age of 4 and the dependent variable applies also to their age of 3 from 7 in range, one must not control for any variable that applies to later circumstances or events in the life cycle and might have been affected by years spent in kindergarten. Otherwise simultaneity concerns arise. For example, one should not control for an indicator that describes whether the child had to repeat any school year since this could be also an outcome variable, similar to test scores. Analogously, one should not keep extracurricular schooling activities, like repetition or workgroup-activities fixed. Participation could be systematically related to kindergarten attendance, either in positive or negative way.

To sum up, variables related both to kindergarten attendance and later test scores should be controlled for which are already present *before* the child's kindergarten attendance becomes relevant or is unlikely being affected by it at later life cycle stages (during school). I include the following rich set of control variables in the analysis: the municipality's characteristics the family lives at (*region and type*), age of the child when she started school, the composition of the household (*the number of the members of the household, whether the child lives together with her mother and father versus stepmother and stepfather*), the employment history of the father, the highest educational attainment of the parents, indicators about the family's

[19] See e.g. Wooldridge (1999, pp. 291).

wealth (*whether the child is considered disadvantaged*[20] *and has to work regularly on the field or in the garden)* and cultural goods (e.g. *access to internet, the number of books at home*). Table3 and Table4 in the Appendix shows detailed definition of and descriptive statistics for the aforementioned variables.

[20] According to Act No. LXXIX of 1993 on Public Education (121 § (1)) a child is disadvantaged if her family is eligible for childprotection support based on her social circumstances.

Chapter 4: Descriptive Statistics

In this Chapter I show the main descriptive statistics. First, I present statistics about the relationship of interest between time spent in kindergarten and student achievement by highest parental educational attainment. Then I turn to the instrument and pay particular attention to its spatial distribution in Hungary to see whether the instrument can be assumed to be distributed randomly. I also present descriptive evidence about the reduced form relationship between test score and scarcity of kindergarten teachers. After that I assess the validity of the instrument. I show the 1st stage relationship between kindergarten attendance and scarcity of kindergarten teachers. Additionally, I show how the instrument is related to other signs of inferior kindergarten service quality. Finally, I examine whether the instrument is correlated with other factors than kindergarten attendance that affect student achievement.

I conclude that the scarcity of kindergarten teachers seems to be non-randomly distributed in space, concentrated in poorer counties. It seems to be systematically related to kindergarten service quality through class size, teacher quality and public kindergarten expenditures per enrolled child. The instrument is systematically related to time spent in kindergarten. Unfortunately, it also seems to be related to student achievement through primary school quality in the dimensions of primary school teacher quality and public general educational expenditures.

4.1. Statistics about time spent in kindergarten and student achievement

Table1 shows the distribution of children by their years spent in kindergarten and their parents' education. 27% of children whose parents finished at most their 8th grade spent less than 3 years in kindergarten. Among children whose parents finished at most vocational, secondary and tertiary education this figure is 12.2%, 7.4% and 6.3%, respectively. The distribution of students whose parents obtained at most a vocational degree is the most similar to the average distribution. Compared to them, the fraction of more disadvantaged students is higher in categories of less kindergarten attendance and lower in categories of more kindergarten attendance. In case of less disadvantaged students the opposite prevails. Thus, there is a monotonic relationship between kindergarten attendance and parental education. Children with less educated parents have spent systematically fewer years in kindergarten.

Table1. Percentage of children by years spent in kindergarten and parental education

	Years spent in kindergarten, 6th grade							
Parental education	**0**	**0,5**	**1**	**1,5**	**2**	**3**	**4**	**Total**
0-8th grade	1,1%	1,0%	7,1%	2,8%	15,3%	45,0%	27,7%	100% (N=8230)
vocational	0,2%	0,2%	2,5%	1,2%	8,1%	47,4%	40,4%	100% (N=24214)
secondary	0,1%	0,2%	1,3%	0,8%	4,9%	47,3%	45,3%	100% (N=26316)
tertiary	0,2%	0,3%	1,1%	0,8%	3,9%	49,0%	44,7%	100% (N=21042)
total	**0,3%**	**0,3%**	**2,2%**	**1,1%**	**6,7%**	**47,6%**	**41,9%**	**100% (N=79802)**

Notes: Own calculations. Data source: HNABC 2008. Children with special needs excluded.

Table2 shows average test scores for children with different family backgrounds and kindergarten attendance. In every category of kindergarten attendance children stemming from more disadvantaged families perform systematically worse. Thus, keeping time spent in kindergarten fixed, there is a positive and monotonic relationship between family background and student achievement. Keeping parental education fixed, children who spent more time in kindergarten perform on average better, although the relationship is not completely monotonic.

The last column of **Table2** shows the difference between average test scores measured in the percentage of standard deviation between children who spent less than 3 years in kindergarten and those who spent 3 or more years, broken by parental education. For children with least educated parents this difference is 22% in mathematics and 29% in reading (measured in standard deviations). For children whose parents obtained at most a vocational degree this difference is 13% in mathematics and 15% in reading. For those whose parents finished secondary school, this difference is 8% in mathematics and 10% in reading, while for children with tertiary educated parents it is 11% for both types of score. Thus, the differences decline with parental education. In general the differences are larger in case of the reading score than in case of mathematics score.

Consequently, the relationship of interest between time spent in kindergarten and test score is positive. Additionally, the difference between the average test scores of children with less and more time spent in kindergarten is the largest for disadvantaged children. This might be a sign that children from less privileged families might gain more from more time spent in kindergarten, although the aforementioned figures do not have any causal interpretation.

Table2. Average test scores by years spent in kindergarten and parental education

Average mathematics score by Years spent in kindergarten, 6th grade								
Parental education	**0**	**0,5**	**1**	**1,5**	**2**	**3**	**4**	**Diff [3/more years – less than 3 years]**
0-8th grade	-1,189	-1,232	-0,935	-1,162	-0,985	-0,804	-0,748	22%
vocational	-0,469	-0,487	-0,419	-0,487	-0,388	-0,314	-0,237	13%
secondary	0,298	0,077	0,033	0,058	0,108	0,161	0,224	10%
tertiary	0,321	0,382	0,424	0,439	0,562	0,601	0,624	11%
Average reading score by Years spent in kindergarten, 6th grade								
Parental education	**0**	**0,5**	**1**	**1,5**	**2**	**3**	**4**	**Diff [3/more years – less than 3 years]**
0-8th grade	-1,246	-1,305	-1,104	-1,284	-1,051	-0,852	-0,754	29%
vocational	-0,554	-0,520	-0,448	-0,465	-0,383	-0,297	-0,207	15%
secondary	0,240	0,075	0,128	0,076	0,124	0,175	0,238	8%
tertiary	0,475	0,396	0,489	0,425	0,556	0,592	0,616	11%

Notes: Own calculations. Data source: HNABC 2008. Children with special needs excluded. Average standardized test scores with mean 0 and standard deviation 1 are shown in the table. The differences in the last column are measured in percentage of one standard deviation.

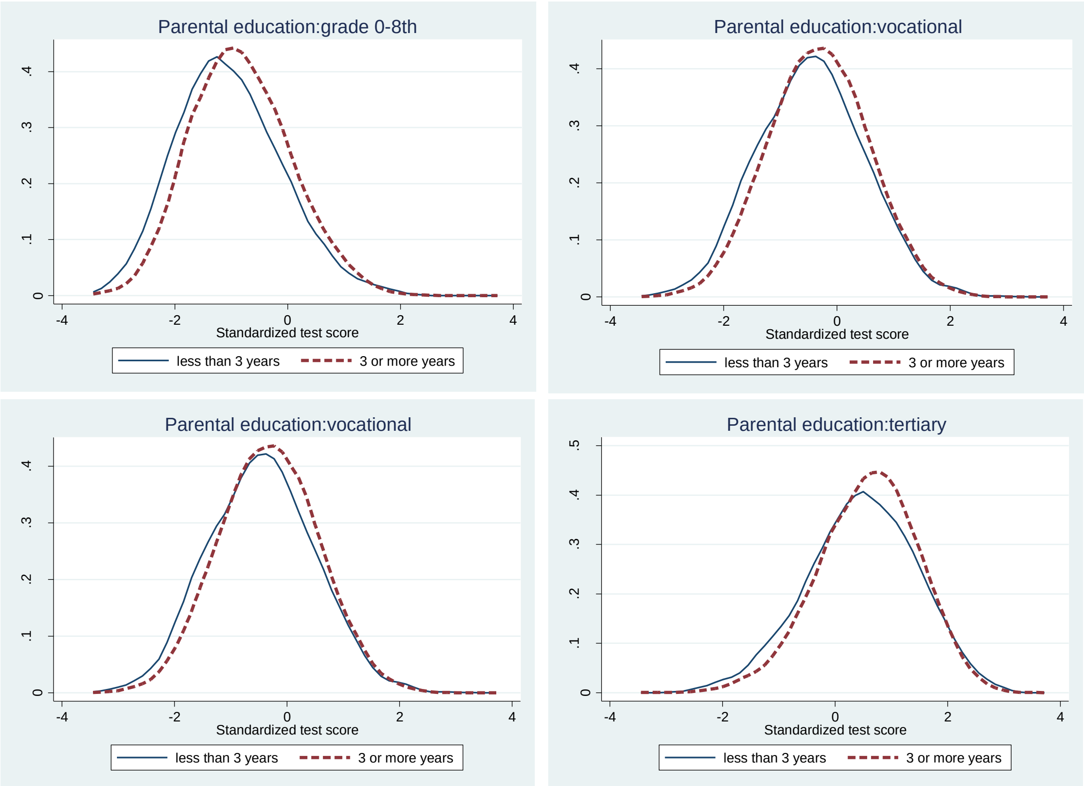
Figure1 shows the probability distribution functions for students who spent less than 3 years in kindergarten and for those who spent 3 or more years, broken by parental education and type of score. When the distribution functions overlap, there is no difference between the student achievements of those who attended kindergarten for different amount of time. The overall pattern is similar for each category of parental background, grade and score. In the range of lower student achievement the probability distribution function for children who spent less time in kindergarten lies above the probability distribution function for children who spent more time in kindergarten. In the range of higher student achievement the opposite prevails.

Differences are the largest for those with the least educated parents, irrespective or score. For students from more privileged families the pattern is mixed. There is almost a perfect overlap in distribution of student achievement for students whose parents finished secondary or tertiary education in case of the reading, while case of the mathematics score the difference is substantial. For students with parents who finished vocational education the degree of overlapping is in between.

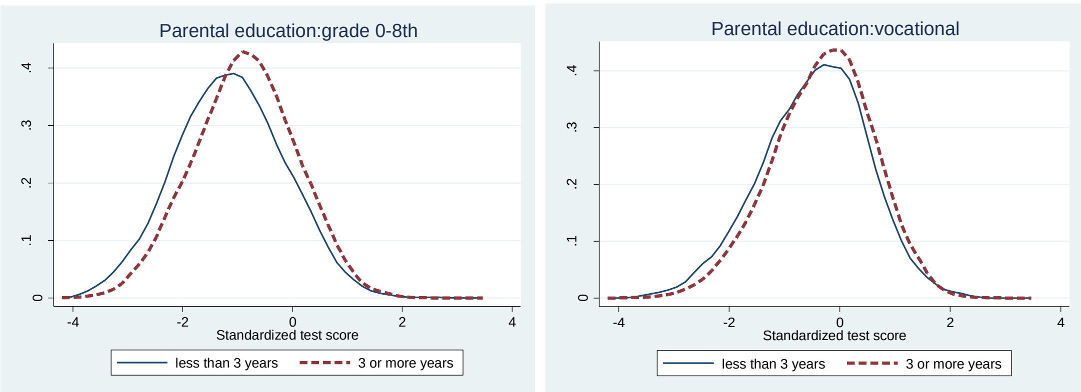
Hence, the difference between the average test scores of children with less and more time spent in kindergarten is a consequence of the differences in the entire distribution of test scores of children with less and more time spent in kindergarten. The differences in the entire distribution of test scores are the largest for the most disadvantaged children.

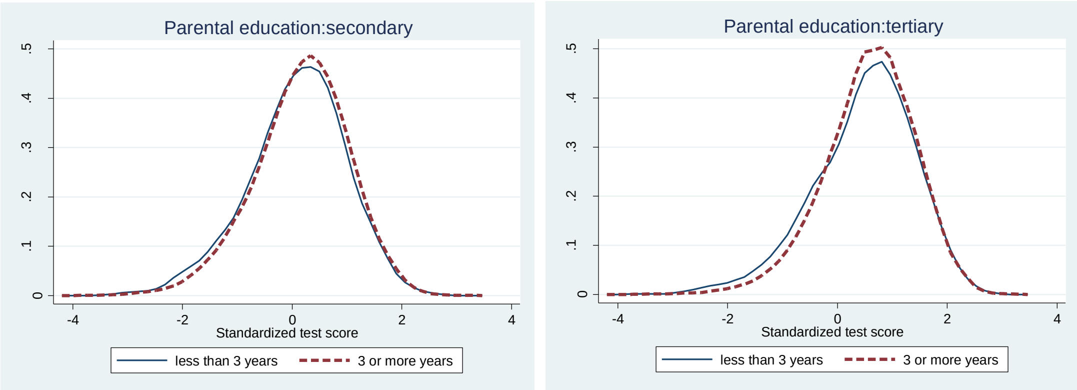
Figure1. Density functions of test scores by years spent in kindergarten and parental education

Mathematics, 6th grade



Reading, 6th grade



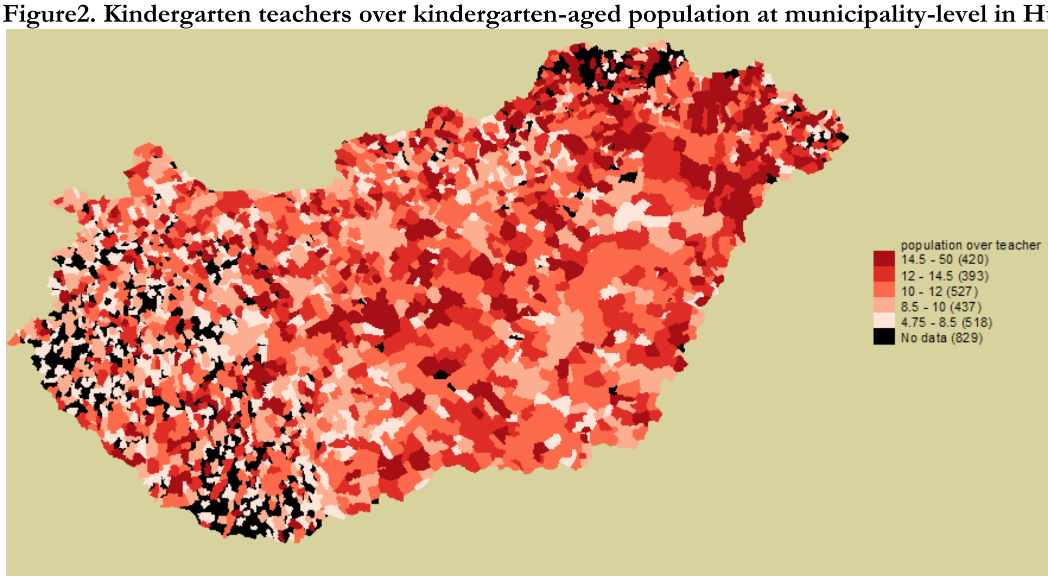


Notes: Own calculations. Data source: HNABC 2008 . Children with special needs excluded. The distribution of average standardized test scores with mean 0 and standard deviation 1 are shown.

4.2. SPATIAL DISTRIBUTION OF THE INSTRUMENT AND THE REDUCED FORM RELATIONSHIP

The instrument is the number of kindergarten teachers relative to the kindergarten-aged population in the particular municipality the family lived at when the child was 4 years old.

Figure2. Kindergarten teachers over kindergarten-aged population at municipality-level in Hungary, 1999



Notes: Own calculations. Data sources: KIR-Stat 1999, DEM 1999. Kindergarten-aged population is assumed to be the 3-5 aged children in the municipality. "No data" refers to no kindergarten service provided in the municipality. Number of observations: 3124.

Figure2 shows how the relative number of kindergarten teachers was distributed across Hungarian municipalities in 1999, when the majority of 6th graders in 2008 were 4 years old. Municipalities colored black did not provide any kindergarten service. Municipalities with more intensive dark colors suffered from increasingly serious scarcity of kindergarten teachers. Hence, darker colors indicate that one teacher had to take care of more children (provided that every kindergarten-aged inhabitant child attended kindergarten in that particular municipality).

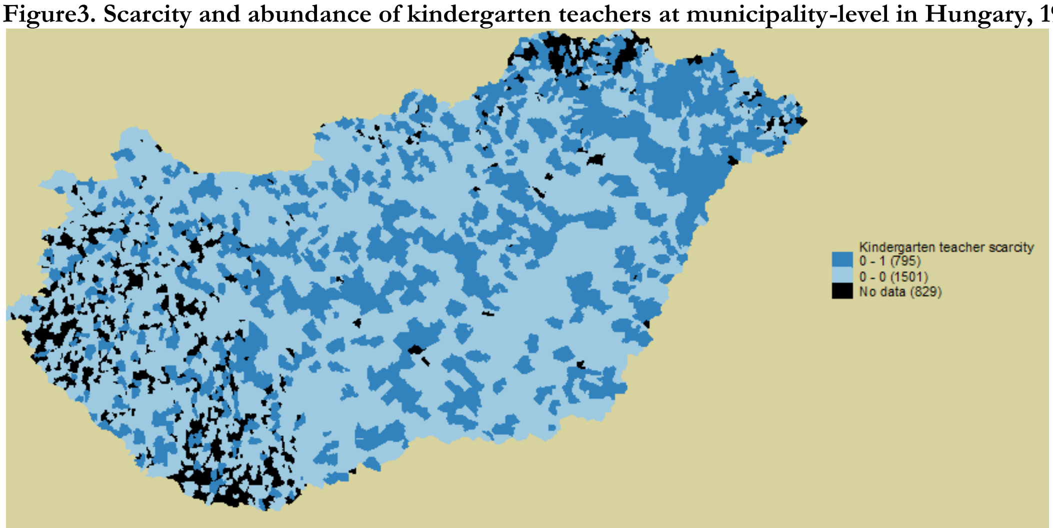
The lack of any kindergarten service was the most characteristic for the southern-western part of Transdanubia (predominantly Vas, Zala, Somogy, Baranya counties) and the northern-eastern part of Hungary (Borsod-Abaúj-Zemplén and Szabolcs-Szatmár-Bereg counties). While there were relatively the most municipalities with no kindergarten services in the southern-western part of Transdanubia, the picture was mixed whether a municipality there offering such service suffered from scarcity or not. In the northern-eastern part of Hungary the picture was clearer: municipalities were on average less likely to offer any kindergarten service compared to the remainder part of Hungary. However, even if they did, the number of teachers seemed to be insufficient compared to the kindergarten-aged population living in the municipality.[21]

Hence, the relative number of kindergarten teachers was not randomly distributed in space in 1999 in Hungary. In the southern-western and northern-eastern part of Hungary and in Central Hungary kindergarten teachers had to take care of systematically more children (provided that every kindergarten-aged inhabitant child attended kindergarten in that particular municipality).

As mentioned in Chapter 3, I defined for each individual a dummy variable of scarcity that equals 1 if the observation lived in a municipality at her age of 4 where the teacher over kindergarten-aged population ratio happened to be at the lowest sextile and 0 otherwise. Thus, scarcity of kindergarten teachers affects approximately 16% of students in the HNABC 2008 data base. 19% of these students lived in Central Hungary, 8.9% in Central Transdanubia, 6.6% in Western Transdanubia, 8.5% in Southern Transdanubia, 21.3% in Northern Hungary, 26.3% in Northern Great Plain and 10% in Southern Great Plain at their age of 4. I selected the municipalities these children lived at their age of 4 and plotted them in 1999.

Figure3 shows how scarcity of kindergarten teachers was distributed across Hungarian municipalities in 1999. Darker blue color indicates scarcity, while brighter blue color indicates abundance. Scarcity was concentrated in the southern-western and northern-eastern part of Hungary.

Figure3. Scarcity and abundance of kindergarten teachers at municipality-level in Hungary, 1999



Notes: Own calculations. Data sources: HNABC 2008, KIR-Stat 1996-2001, DEM 1996-2001. Scarcity of kindergarten teachers equals one if kindergarten teachers over the 3-5 aged population was lower than the 16. percentile at the child's age of 4 in the HNABC data base. "No data" refers to no kindergarten service provided in the municipality. Number of observations: 3124.

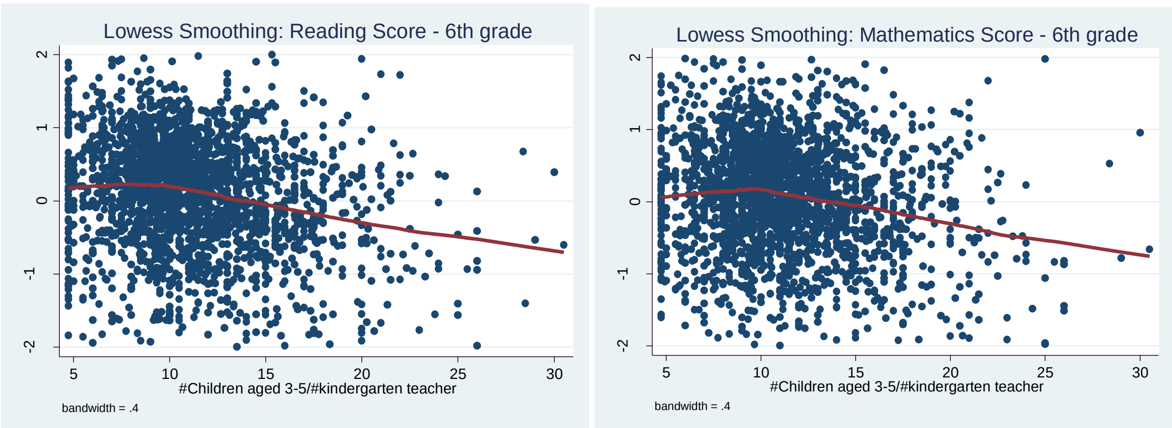
[21] Taking a look at the spatial distribution of mean kindergarten attendance, children living in Northern Great Plain and Northern Hungary attended kindergarten for systematically fewer years on average, while there is no such pattern for Southern or Western Transdanubia. See Table5 in the Appendix for further details.

Consequently, the instrument was not randomly distributed in space, but was concentrated in counties of Hungary which are on average substantially poorer than other counties in Hungary[22]. If the assumptions about the systematic relationship between family background and individual-specific components of marginal benefits, marginal costs and ability are safe, then the instrument does not satisfy the exogeneity assumption in Equation (21). From the discussion of the endogeneity channel in Section 2.3. it follows that the systematic relationship between the instrument and individual-specific parameters very likely introduces an upward bias in the IV-estimates.

Figure4 shows the reduced form relationship between the average test score of the municipality and the kindergarten-aged population over teacher ratio. The lowess smoothing figures show that up to 10 kindergarten-aged inhabitant children a teacher has to take care of, there is no or only a slightly positive relationship between test scores and teacher scarcity. Above 10 children one obtains a negative relationship. This pattern remains irrespective of scores.

Consequently, the reduced form relationship seems to be significant. Especially, if one takes into consideration that the cutoff-point for the instrument in binary form is in the range of relative number of teachers where the relationship is unambiguously negative. If the instrument is valid, then the only reason why test scores vary with the instrument should be that time spent in kindergarten varies with the instrument.

Figure4. Average test score in 2008 and kindergarten-aged population over teacher ratio in 1999 at municipality-level



Notes: Own calculations. Data source: HNABC 2008, KIR-Stat 1996-2001 and DEM 1996-2001. Children with special needs do not count into the average score of the municipality. Average standardized test scores with mean 0 and standard deviation 1 are shown. Children over kindergarten teacher ratio is censored from below (3%). Number of observations: 2199. The following municipalities are excluded:
- average test score is outside the (-2,2) standard deviation range
- provides no kindergarten service, thus the number of kindergarten teachers are zero
- one kindergarten teachers has to take care of more than 30 children on average.

To be valid, the chosen instrument has to satisfy two crucial assumptions. First, it has to be significantly related to the time spent in kindergarten. Children who lived in municipalities at their age of 4

[22] For instance in 1999, income per capita was 192 thousand forints in Szabolcs-Szatmár-Bereg county, 241 thousand forints in Borsod-Abaúj-Zemplén county, 239 thousand forints in Somogy county, 266 thousand forints in Baranya county and 295 thousand forints in Zala county. The national average was 309 thousand forints measured in 1999-forints. *(Data source: Hungarian Regional Database of the Hungarian Academy of Sciences, http://www.eroforrasterkep.hu/index.php?setLang=en).*

where there happened to be a shortage of kindergarten teachers should have spent systematically less time in kindergarten. Second, serious kindergarten teacher scarcity is assumed to have no effect on test score other than operating through kindergarten attendance. In the next subsection I assess separately whether these assumptions can be justified.

4.3. Assessing the validity of the instrument

4.3.1. The 1st stage relationship and the mechanism of interest

In this subsection first I illustrate the 1st stage relationship between time spent in kindergarten and scarcity in kindergarten teachers. One of the requirements for having a valid instrument is to have a strong 1st stage relationship. As Bound et al. (1995) emphasize, the use of weak instruments can lead to large inconsistencies in the IV estimates, even if there is a weak relationship between unobservables in the relationship of interest and the instrument.

Then I show that municipalities with scarcity of kindergarten teachers provided on average kindergarten service with lower quality according to other measures, like class size and public expenditures, but not according to kindergarten teacher quality. This is important to assess whether the instrument affects individual-specific benefit and cost terms through inferior kindergarten service quality. The potential effect is part of the mechanism as I discussed in Section 2.3.

Column1 in **Table3** shows that children who attended kindergarten for 3 or 4 years happened to live in municipalities at their age of 4 where a kindergarten teacher had to take care of on average less children (provided that every kindergarten-aged inhabitant child would have attended kindergarten in that particular municipality). These children lived in municipalities at their age of 4 where, on average, 11 children could have been taken care of by one teacher. Children who attended kindergarten for fewer years lived in municipalities where, on average, 13 children could have been taken care of by one teacher. Column2 in **Table3** shows that the ratio of children who lived in a municipality at age 4 where there was scarcity of kindergarten teachers is twice among children who attended kindergarten for less than 3 years. Consequently, access to kindergarten service seems to be systematically related to attendance.

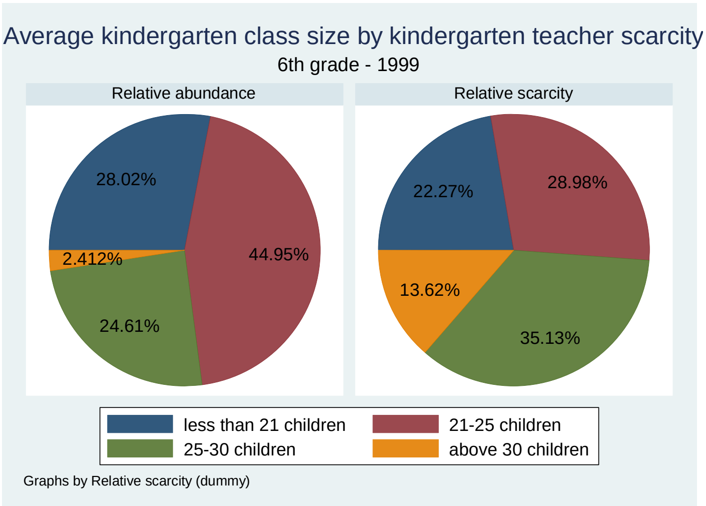
Table3. Kindergarten teacher scarcity measures at municipality-level by years spent in kindergarten

6th graders years spent in kindergarten	**kindergarten-aged population over teacher**	**ratio of children lived in municipality with scarcity of teachers**	**N**
0 year	12,62	33,02%	216
0,5 year	13,24	32,55%	257
1 year	13,77	39,92%	1 836
1,5 years	12,15	27,41%	937
2 years	12,60	33,00%	5 611
3 years	10,90	16,36%	39 467
4 years	10,84	12,83%	35 036

Notes: Own calculations. Data sources: HNABC 2008 KIR-Stat 1996-2001, DEM 1996-2001. Children with special needs excluded. Scarcity of kindergarten teacher is one if kindergarten teachers over the 3-5 aged population was lower than the 16. percentile at the child's age of 4 in the HNABC data base. Kindergarten teacher over population ratio is censored from above (3%).

Figure5 shows how large was the fraction of the average number of particular kindergarten classes[23] in 1999 in municipalities in which individuals in the HNABC 2008 database experienced scarcity *versus* abundance of kindergarten teachers at their age of 4. In 1999 it can be seen that in municipalities experiencing shortage, the ratio of the average kindergarten classes with above 30 children was on average 6-times, the ratio of classes with 25-30 children was on average 43% higher than in municipalities that did not suffer from scarcity. On the contrary, the ratio of kindergarten classes with moderate number (21-25) of children was approx. 35% lower in the aforementioned municipalities. Hence, children who attended kindergarten in municipalities experiencing shortage of kindergarten teachers had to suffer from some kind of crowding-effect. This suggests that besides the relatively low number of kindergarten teachers, in these municipalities the size of the classes was not adjusted properly to avoid crowdedness by reducing the number of admitted children.

Figure5. Distribution of the average number of kindergarten classes with different size, by scarcity of kindergarten teachers, 1999

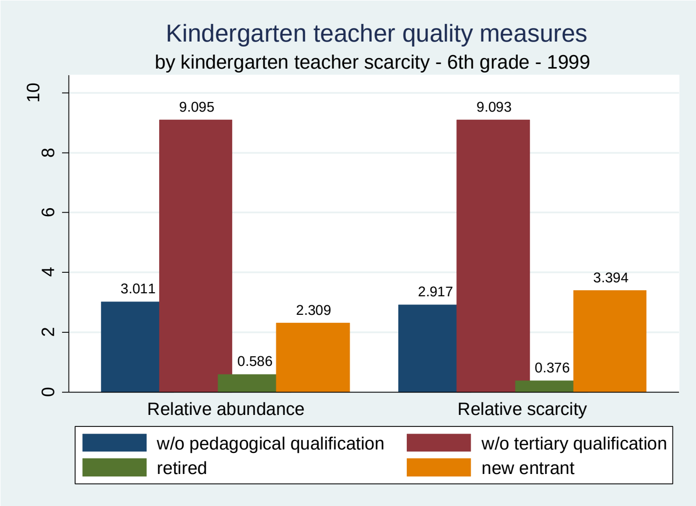


Notes: Own calculations. Data sources: HNABC, KIR-Stat 1996-2001, DEM 1996-2001. Scarcity of kindergarten teachers equals one if kindergarten teachers over the 3-5 aged population was lower than the 16. percentile at the child's age of 4 in the HNABC data base. Number of observation is 3124. Number of municipalities with scarcity of kindergarten teachers is 1501.

Figure6 shows the average municipality-level ratios of kindergarten teachers with characteristics that proxy for lower teacher quality in 1999, broken down by scarcity of kindergarten teachers. The considered groups are the following: kindergarten teachers without pedagogical qualification, without tertiary qualification, retired and new entrant kindergarten teachers. There was almost no difference between municipalities in case of the ratio of teachers without pedagogical qualification, without tertiary qualification or retired teachers. At the same time, the ratio of new entrant teachers was by 45% higher in municipalities that experienced kindergarten teacher scarcity in 1999. Taking into consideration the evidence on new entrant teachers' potential low quality of teaching in their first year of tenure (Hanushek et al., 2002), this indicates that municipalities that experienced serious shortage of kindergarten teachers in 1999 employed kindergarten teachers of lower quality.

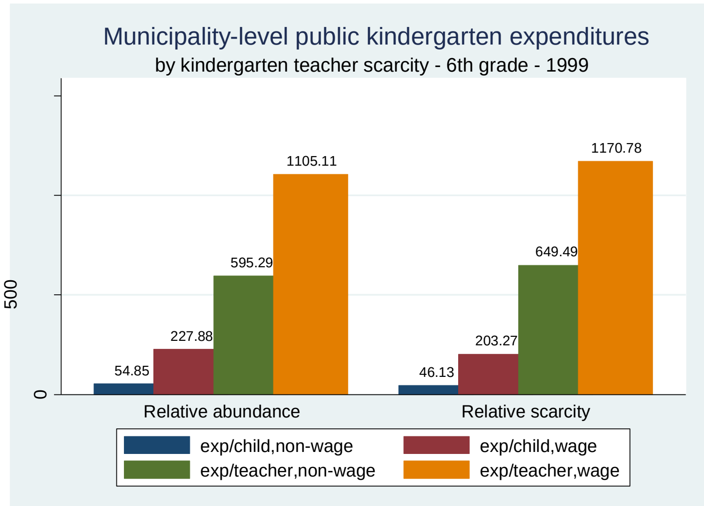
[23] I.e. classes with less than 21 children, with 21-25 children, with 26-30 children and above 30 children.

Figure6. Kindergarten teacher quality measures, by scarcity of kindergarten teachers, 1999



Notes: Own calculations. Data sources: HNABC, KIR-Stat 1996-2001, DEM 1996-2001. Scarcity of kindergarten teachers equals one if kindergarten teachers over the 3-5 aged population was lower than the 16. percentile at the child's age of 4 in the HNABC data base. Number of observation is 3124. Number of municipalities with scarcity of kindergarten teachers is 1501.

Figure7. Municipality-level public kindergarten expenditures, by scarcity of kindergarten teachers, 1999



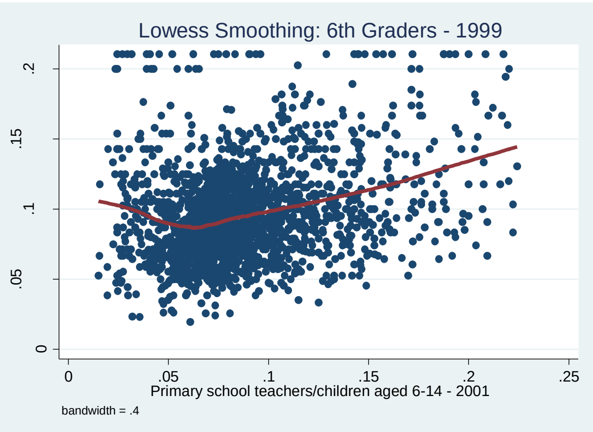
Notes: Own calculations. Data sources: HNABC, KIR-Stat 1996-2001, DEM 1996-2001, MPE 1999. Scarcity of kindergarten teachers equals one if kindergarten teachers over the 3-5 aged population was lower than the 16. percentile at the child's age of 4 in the HNABC data base. Number of observation is 3124. Number of municipalities with scarcity of kindergarten teachers is 1501. Amounts are measured in 1999-forints.

Figure7 shows average kindergarten expenditures per teacher and per enrolled child in 1999 in municipalities in which individuals in the HNABC 2008 database experienced scarcity *versus* abundance of kindergarten teachers at their age of 4. The average non-wage expenditures per teacher were 9% (54 thousand forints) higher in municipalities where the relative number of teachers was particularly low in 1999. Municipalities experiencing scarcity in 1999 spent on average 6% (65 thousand forints) more on wage per teacher. The average wage expenditures per enrolled child were 12% (25 thousand forints), while the average non-wage expenditures per enrolled child were 19% (9 thousand forints) higher in municipalities with serious teacher scarcity in 1999. The overall picture is mixed: while expenditures per teachers were higher, expenditures per children were lower in the municipalities that experienced scarcity of kindergarten teachers. This implies that these municipalities provided kindergarten service of lower quality in the dimension of expenditures per enrolled child[24].

4.3.2. The relationship between scarcity of kindergarten teachers and primary school quality

One of the concerns regarding the validity of the instrument is that it could lead to lower test scores independently of the relative number of kindergarten teachers. This might happen if municipalities experiencing a shortage of kindergarten teachers also provide inferior schooling quality in primary school. In this subchapter I provide descriptive evidence about primary school quality in the municipality at the child's age of 6 (thus, at school-starting age) for children who lived in municipalities with kindergarten teacher scarcity at their age of 4. The main result is that there is some sign that municipalities with serious shortage from kindergarten teachers also provided inferior schooling quality with respect to teacher quality. However, the same does not hold with respect to class size and public expenditure.

Figure8. Kindergarten teachers over kindergarten-aged population in 1999 and primary school teachers over primary school aged population in 2001, at municipality-level



Notes: Own calculations. Data sources: HNABC 2008 KIR-Stat 1999/2001 and DEM 1999/2001. Teacher over children ratio is censored from above (3%). Number of observations 2122. The following municipalities are excluded:
- provides no kindergarten service, thus the number of kindergarten teachers are zero
- primary school teachers over kindergarten aged population is below 0.015 or above 0.225.

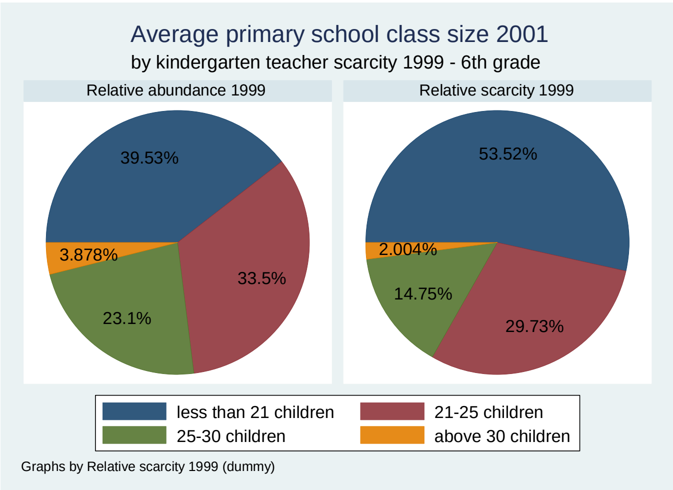
Figure8 shows how kindergarten teachers over kindergarten-aged population at the child's age of 4 and primary school teachers over primary school aged population at the child's age of 6 were related at municipality-level. Below a cutoff-value (approx. 0.05 [25]) there was a negative, while above it there was a positive relationship between relative number of kindergarten and primary school teachers. Since for the majority of the municipalities the ratio of primary school teachers and primary school aged population was above this cutoff-value in 1999[26], the observed positive relationship is more characteristic. Hence, municipalities experiencing a relative shortage of kindergarten teachers in 1999 were likely to experience

[24] Although teacher quality has been found to play more important role in determining school quality than school expenditures (Hanushek, 2002 or Rivers and Sanders, 2002).
[25] This cutoff-value represents that one primary school teacher has to teach on average 20 students, if all kindergarten-aged child goes to primary school in the particular municipality.
[26] In 1999, 12% of the municipalities had this ratio of primary school teachers and primary school aged below this cutoff-value.

also a relative shortage of primary school teachers two years later, when the majority of the kindergarten-aged children started primary school.

Another dimension of school quality is the average class size. **Figure9** shows that in municipalities that experienced shortage of kindergarten teachers in 1999, the ratio of the average number of primary school classes with below 21 children was by 36% higher in 2001, while the ratio the average number of primary school classes with 25-30 children was by 35% lower in 2001, than in municipalities that did not suffer from teacher scarcity in 1999. The ratio of the average number of classes with 21-25 and above 30 students in 2001 is almost identical in the two types of settlements. Thus there is no sign that children who lived in municipalities with kindergarten teacher scarcity at their age of 4 would have suffered from some kind of crowding-effect when becoming a primary school student.

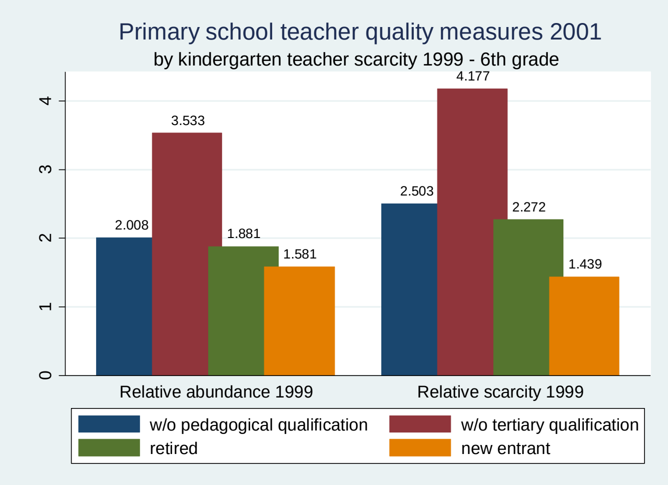
Figure9. Distribution of the average number of primary school classes with different size, by scarcity of kindergarten teachers, 1999



Notes: Own calculations. Data sources: HNABC, KIR-Stat 1996-2001, DEM 1996-2001. Scarcity of kindergarten teachers equals one if kindergarten teachers over the 3-5 aged population was lower than the 16. percentile at the child's age of 4 in the HNABC data base. Number of observation is 3124. Number of municipalities with scarcity of kindergarten teachers is 1501.

Figure10 shows how the ratio of particular type of primary school teachers differed in municipalities in 2001that experienced kindergarten teacher scarcity *versus* abundance two years before. In spite the low absolute level of the ratio of teachers of lower quality, municipalities with scarcity of kindergarten teachers in 1999 employed on average by 25% more full-time teachers without pedagogical qualifications in 2001 than municipalities experiencing abundance. Municipalities with scarcity employed by 18% more full-time teachers without tertiary qualification and by 21% more retired pedagogues in the same year. There was practically no difference between the ratios of new entrants employed in the two types of municipalities. This indicates that children who lived in municipalities with kindergarten teacher scarcity at their age of 4 were likely to have primary school teachers of lower quality when becoming primary school eligible two years later.

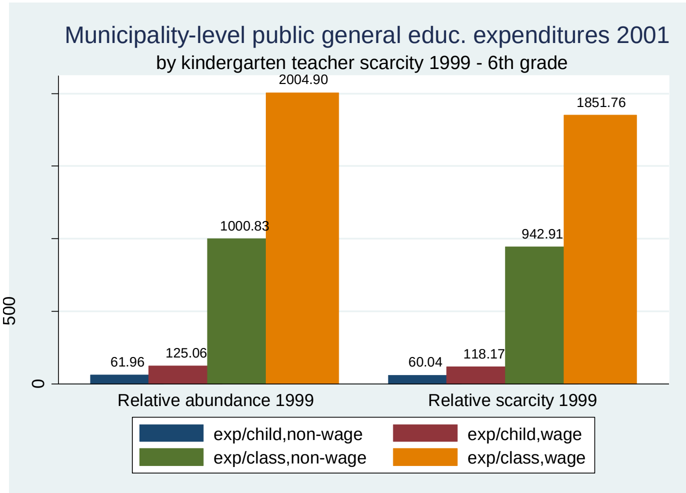
Figure10. Primary school teacher quality measures, by kindergarten teacher scarcity, 1999



Notes: Own calculations. Data sources: HNABC 2008 and KIR-Stat 1997, 1999, 2001. Serious kindergarten teacher scarcity equals one if kindergarten teachers over the 3-5 aged population was lower than the 16. percentile at the child's age of 4 in the HNABC data base. Number of municipalities with serious kindergarten teacher scarcity is N_{1999}*=1501,*N_{1997}*=1531.*

Figure11 shows average municipality-level public general educational expenditures per enrolled child and per class in 2001 broken by scarcity of kindergarten teachers in 1999. There were practically no differences between the average non-wage expenditures per enrolled child. Municipalities experiencing scarcity of kindergarten teachers in 1999 spent on average by 6% (7 thousand forints) less on wage per child in 2001. The average wage expenditures per class were by 8% (153 thousand forints), while the average non-wage expenditures per class were by 6,4% (58 thousand forints) lower in 2001 in municipalities with scarcity in kindergarten teachers in 1999, than in municipalities with abundance. Overall, there is a sign for a systematic relationship between the relative number of kindergarten teachers and public general educational expenditures at municipality-level.

Figure11. Municipality-level public general educational expenditures in 2001, by scarcity in kindergarten teacher scarcity in 1999



Notes: Own calculations. Data sources: HNABC, KIR-Stat 1996-2001, DEM 1996-2001, MPE 2001. Scarcity of kindergarten teachers equals one if kindergarten teachers over the 3-5 aged population was lower than the 16. percentile at the child's age of 4 in the HNABC data base. Number of observation is 3124. Number of municipalities with scarcity of kindergarten teachers is 1501.Amounts are measured in 2001-forints. General education expenditures involve primary school and general secondary school expenditures, but no vocational educational expenditures.

Hence, I find that students, who lived in municipalities with scarcity of kindergarten teachers when they were 4 years old, experienced systematically lower teacher quality and public general educational expenditures two years later in primary school than those who lived in municipalities with abundance of kindergarten teachers when they were 4 years old. The same does not apply to class size that proxy school quality. Thus, even if exogeneity in the reduced form equation can be assumed, it seems to be false that the only channel through the instrument is related to test scores is through the number of years spent in kindergarten. As a consequence, I expect the IV-estimates to be upward biased due to the negative correlation between scarcity of kindergarten teachers and primary school quality on the one hand and the negative correlation between scarcity and time spent in kindergarten on the other.

The statistics so far did not have a causal interpretation. Now I turn to the regression results.

Chapter 5: Regression Results

In this Chapter first I show the result of the OLS and IV estimates. The instrument is a dummy variable indicating whether the kindergarten teacher over kindergarten-aged (3-5 year aged) population ratio was below the 16. percentile (or analogously, in the lowest sextile) in the municipality where the child lived at her age of 4. I present robustness checks with a slight modification of the instrument. The robustness checks look at whether the results are sensitive to changing the number of kindergarten teachers to number of kindergarten places or the reference 3-5 year aged cohorts to 3-6 year aged cohorts. Finally I show the OLS and IV results by parental education to explore any heterogeneous effects.

All presented regression coefficients are from estimations in which I control for all the variables in Table3 in the Appendix. These control variables are the following: the municipality's characteristics the family lives at (*region and type*), age of the child when she started school, the composition of the household (*the number of the members of the household, whether the child lives together with her mother and father versus stepmother and stepfather*), the employment history of the father, the highest educational attainment of the parents, indicators about the family's wealth (*whether the child is considered disadvantageous*[27] *and has to work regularly on the field or in the garden)* and cultural goods (e.g. *access to internet, the number of books at home*).

5.1. Regression results for the whole sample

Table4 shows the OLS and IV estimates of the regression of years spent in kindergarten on mathematics and reading test score for the Hungarian 6th graders in 2008.

Table4. OLS and IV estimates of the effect of time spent in kindergarten on test score, IV 1st stage included

		mathematics	reading
OLS	Years spent in kindergarten	0.0613***	0.0594***
		[0.004]	[0.004]
IV	Years spent in kindergarten	0.1844***	0.2153***
		[0.044]	[0.044]
	1st stage	Years spent in kindergarten	
	Teachers / children aged 3-5 < 16. pctile	-0.2084***	
		[0.008]	

*Notes: Own calculations. Data source: HNABC 2008, KIR-Stat 1996-2001, DEM 1996-2001. Source of the figures: Table5, Table7, Table8 in the Appendix, Column4 and Column8. Children with special needs excluded. All control variables listed in Table3 in the Appendix included. Scores are measured in standard deviations. Robust standard errors in parentheses. Significance levels: *** 1%, ** 5%, * 10%. Instrument: Scarcity of kindergarten teachers that equals one if kindergarten teacher over the 3-5 aged population ratio in the municipality at the child's age of 4 is lower than the 16. percentile in the HNABC data base.*

First consider the OLS estimates. Keeping parental education, school-starting age, poverty and wealth indicators, composition and spatial characteristics of the household fixed, 6th graders who attended kindergarten for one additional year achieved on average by 6% higher score in both mathematics and reading in 2008. The effects are significant at 1% level. Table6 in the Appendix shows that the simple regression estimates are 16.4% and 17.6% in case of mathematics and reading score, respectively. 59% of

[27] According to Act No. LXXIX of 1993 on Public Education (121 § (1)) a child is disadvantaged if her family is eligible for childprotection support based on her social circumstances.

the effect is transmitted through parental education in case of mathematics score and 57% is transmitted through parental education in case of reading score. Overall, 63-66% of the total effect is mediated through the control variables.

The IV estimates are substantially higher. Keeping parental education, poverty and wealth indicators, composition and spatial characteristics of the household fixed, 6th graders who attended kindergarten for one additional year achieved by 18.4% higher score in mathematics and by 21.5% higher score in reading. Both effects are significant at 1% level. (Note that the IV standard errors are 11-times larger than the OLS standard errors.) The effect on reading score is by 16% larger than the effect on mathematics score. From Table7 in the Appendix it can be seen that the simple IV estimates are huge: 117.45% and 127.86% for mathematics and reading score, respectively. Similarly to the OLS estimates, controlling for parental education reduces the effect by 50% in case of mathematics and 55% in case of reading score. Contrary to the OLS-estimates, 83-85% of the total effect is transmitted through the control variables.

The 1st stage estimate shows that the instrument is very strong. According to the last line in Table4, comparing two children with the same parental education, the same social, compositional and spatial characteristics of their family and identical school-starting age, the one who lived in a municipality at her age of 4 where the relative number of kindergarten teachers was extremely low, that child attended kindergarten for 0.21 years less time on average. The t-statistics of the instrument is above 26.

Table8 in the Appendix shows that the 1st stage simple regression coefficient is -0.3023. This means that comparing two children, the one who lived in a municipality at her age of 4 where the relative number of kindergarten teachers was extremely low, attended kindergarten for 0.3 years less time on average*, ceteris paribus*. Thus, all the control variables transmit 30% of the effect, which is substantially lower than in case of the 2nd stage. At the same time, 80% of the effect is mediated through parental education. Hence, parental education has a more important transmission role than in the 2nd stage.

According to Column4 in Table8 in the Appendix, children with least educated parents obtain on average by 0.26 years less kindergarten care than children with parents who obtained a secondary school degree, keeping everything else constant. Children with parents who obtained a vocational degree spent by 0.05 years less years in kindergarten than children with parents who obtained a secondary school degree, *ceteris paribus*. These two effects are significantly different from zero while the length of kindergarten attendance for children with tertiary and secondary educated parents does not differ significantly from each other by controlling for all the controls variables and the scarcity of kindergarten teachers. Thus, children with less educated parents spent on average systematically less time in kindergarten, *ceteris paribus*. This result confirms descriptive ones obtained in **Table1** in Chapter 4.

Table5 presents how the estimated effects differ if I change slightly the instrument.

In the 1st robustness check the instrument is a dummy variable indicating whether the kindergarten places over kindergarten-aged (3-5 year aged) population ratio was below the 16. percentile (or analogously, in the lowest sextile) in the municipality where the child lived at her age of 4. Instrumenting with this variable and keeping parental education, poverty and wealth indicators, composition and spatial characteristics of the household fixed, 6th graders who attended kindergarten for one additional year achieved by 26% higher score in mathematics and by 22% higher score in reading. Thus, the IV-estimates increase by 8 percentage point in mathematics and 1 percentage point in reading when using kindergarten places instead of teachers as a measure of kindergarten service provision.

Table5. OLS and IV estimates of the effect of time spent in kindergarten on test score; various instruments

Panel A	mathematics	reading	Panel B: 1st stage: Years spent in kindergarten
OLS	0.0613***	0.0594***	-
	[0.004]	[0.004]	
IV	0.1844***	0.2153***	-0.2084***
	[0.044]	[0.044]	[0.008]
Robustness Check #1	0.2642***	0.2241***	-0.2503***
	[0.036]	[0.036]	[0.008]
Robustness Check #2	0.1609***	0.2001***	-0.2114***
	[0.043]	[0.043]	[0.008]
Robustness Check #3	0.2321***	0.1946***	-0.2512***
	[0.036]	[0.036]	[0.008]

*Notes: Own calculations. Data source: HNABC 2008, KIR-Stat 1996-2001, DEM 1996-2001. Source of the figures: Table9 and Table10 in the Appendix. Children with special needs excluded. All control variables listed in Table3 in the Appendix included. Scores are measured in standard deviations. Robust standard errors in parentheses. Significance levels: *** 1%, ** 5%, * 10%. Instruments: Robustness Check #1: kindergarten places over kindergarten-aged (3-5 aged) population ratio in the municipality the child lived at age 4 is in the lowest sextile in the HNABC data base. Robustness Check #2: kindergarten teachers over kindergarten-aged (3-6 aged) population ratio in the municipality the child lived at age 4 is in the lowest sextile in the HNABC data base. Robustness Check #3: kindergarten places over kindergarten-aged (3-6 aged) population ratio in the municipality the child lived at age 4 is in the lowest sextile in the HNABC data base.*

In the 2nd robustness check the instrument is a dummy variable indicating whether the kindergarten teacher over kindergarten-aged population ratio was below the 16. percentile (or analogously, in the lowest sextile) in the municipality where the child lived at her age of 4. In this case I define kindergarten-aged children to be between 3 and 6 years old. Instrumenting with this variable and keeping parental education, poverty and wealth indicators, composition and spatial characteristics of the household fixed, 6th graders who attended kindergarten for one additional year achieved by 16% higher score in mathematics and by 20% higher score in reading. The IV-estimates decreases by 2.5 percentage point in mathematics and 1.5 percentage point in reading. Thus, the results are not sensitive to whether I take into account children who stay in kindergarten for a 4th year (although this margin turned out to be important at individual-level, see **Table1**).

In the 3rd robustness check the instrument is a dummy variable indicating whether the kindergarten places over kindergarten-aged population ratio was below the 16. percentile in the municipality where the child lived at her age of 4. In this case I also define kindergarten-aged children to be between 3 and 6 years old. The pattern is similar to what I find in case of kindergarten teachers. If taking into account children who stay in kindergarten for a 4th year, the IV-estimates decrease, but not substantially.

Regarding the 1st stage estimates, the t-statistics increased remarkably from 26 to 31 in case of the 1st robustness check, while increased slightly from 26 to 26.5 in case of the 2nd robustness check. This suggests that the availability of places is more binding for families and using kindergarten places instead of teachers as a measure of kindergarten service provision might lead to a stronger instrument. However, presumably there is less exogeneous variation behind the creation and termination of kindergarten places than behind the hiring or layoff of kindergarten teachers that causes scarcity. Therefore, the larger IV-estimates in case of places are very likely to be more upward biased than the IV-estimates in case of teachers.

To sum up, the IV results seem to be robust to a slight modification in the denominator by extending the kindergarten-aged population from 3-5 aged to 3-6 aged. If I proxy kindergarten service provision by kindergarten places instead of kindergarten teachers, the IV estimates are larger. This confirms the intuition that the number of places in a municipality is very likely to be more endogeneous than the number of kindergarten teachers.

5.2. Regression results by parental education

Table6 shows the OLS and IV estimates by parental education.

Table6. OLS and IV estimates of the effect of time spent in kindergarten on test score, by parental education, IV 1st stage included

Panel A		mathematics			
6th grade	parental education:	grade 0-8	vocational	secondary	tertiary
OLS	**Years spent in kindergarten**	0.0591***	0.0518***	0.0655***	0.0587***
		[0.011]	[0.008]	[0.008]	[0.010]
IV	**Years spent in kindergarten**	0.1388**	0.2169***	0.1195	0.1982
		[0.066]	[0.061]	[0.117]	[0.366]
		reading			
	parental education:	grade 0-8	vocational	secondary	tertiary
OLS	**Years spent in kindergarten**	0.0775***	0.0523***	0.0525***	0.0444***
		[0.011]	[0.008]	[0.008]	[0.009]
IV	**Years spent in kindergarten**	0.1434**	0.2513***	0.0784	0.4392
		[0.069]	[0.062]	[0.116]	[0.366]
Panel B: 1st stage		**Years spent in kindergarten**			
	Teachers / children aged 3-5 < 16.pctile	-0.3472***	-0.2444***	-0.1381***	-0.0628***
		[0.024]	[0.013]	[0.014]	[0.018]

*Notes: Own calculations. Data source: HNABC 2008, KIR-Stat 1996-2001, DEM 1996-2001. Source of the figures: Table11, Table12 and Table13 in the Appendix. Children with special needs excluded. All control variables listed in Table3 in the Appendix included. Scores are measured in standard deviations. Robust standard errors in parentheses. Significance levels: *** 1%, ** 5%, * 10%. Instruments: kindergarten teacher over kindergarten-aged (3-5 aged) population ratio in the municipality the child lived at age 4 is in the lowest sextile in the HNABC data base.*

First consider the OLS estimates and children with least educated parents. Keeping poverty and wealth indicators, composition and spatial characteristics of the household fixed, 6th graders who attended kindergarten for one additional year have on average by 5.9% higher score in mathematics and 7.7% higher score in reading (measured in standard deviation). The same figures for students whose parents who obtained vocational, secondary and tertiary degree are 5.1%, 6.5% and 5.9% in case of mathematics and 5.2%, 5.2% and 4.44% in case of reading. All coefficients are significant at 1% significance level. Thus, in case of reading there is a monotonic relationship between the magnitude of the estimated coefficient and parental education: children from less privileged families benefit more from additional year spent in kindergarten. At the same time, the effects are very close to each other in case of mathematics and there is no monotonic pattern.

Now consider the IV estimates and children with least educated parents. Keeping poverty and wealth indicators, composition and spatial characteristics of the household fixed, 6th graders who attended kindergarten for one additional year have on average by 14% higher score both in mathematics and in reading (measured in standard deviation). 6th graders whose parents finished vocational school have on average by 22% higher score in mathematics and by 25% in reading. The aforementioned coefficients are significantly different from zero at any significance level. There is no significant effect for students whose parents obtained a secondary school or a tertiary degree, either in mathematics or in reading. Consequently, the IV estimates confirm that children from less privileged families benefit more from additional year spent in kindergarten.

Note that the standard errors of the IV estimates are substantially higher in case of students with higher parental education. The 1st stage estimates reveal the causes. I obtain both higher 1st stage effect (in absolute value) and t-statistics in case of children from less peivileged families. Comparing two children whose parents finished at most the 8th grade in primary school, the one who lived in a municipality at her age of 4 where the relative number of kindergarten teachers was extremely low attended kindergarten for 0.35 years less time on average*, ceteris paribus* (t-statistics 14.46). Comparing two children whose parents obtained a vocational degree, the one who lived in a municipality with scarcity of teachers attended kindergarten for 0.24 years less time on average*, ceteris paribus* (t-statistics 18). The effects and t-statistics are 0.14 years and 9.86 for children whose parents finished secondary school and 0.06 years and 3.5 for children whose parents obtained a tertiary degree. This suggests that less privileged families are more constrained in case of scarcity of kindergarten service provision. Their stronger 1st stage relationship translates into lower standard deviations[28].

To sum up, for disadvantaged families the cost of kindergarten attendance arising due to scarcity of kindergarten service provision seems to be higher than for privileged families. At the same time, the results suggest that less privileged children benefit more from an additional year spent in kindergarten.

One caveat: recall the LATE interpretation discussed in Section 2.3. LATE is by definition the average causal effect of treatment for individuals whose treatment status if affected by the change in the instrument. Recall also, that the constrained individuals may or may not be representative of the entire population, and it is impossible to identify them. Also, the average causal response is a weighted average of the individual-specific slopes along a response function, where the weights are proportional to the number of affected individuals. Thus, LATE is the treatment effect identified for whom this constraint of scarcity of kindergarten service is most binding. As the 1st stage estimates reveal, LATE is the treatment effect identified mainly for children from the less privileged families. Consequently, it might happen that the effect for privileged children is also significantly positive, but they are affected to a less extent, thus the effect is identified for them to a less extent.

Finally, it might be useful to review the potential causes of the substantially higher IV estimates (compared to the OLS estimates which identifies average effects).

[28] The asymptotic variance of the IV estimates has the squared correlation coefficient between the instrument and the endogenous variable in its denominator, thus the stronger the 1st stage, the lower the asymptotic variance, *ceteris paribus* (see e.g. Wooldridge, 1999, pp. 466).

First, the instrument is non-randomly distributed in space, but concentrated in counties of Hungary which are on average substantially poorer than other counties in Hungary. From the discussion of the endogeneity channel in Section 2.3. it follows that the systematic relationship between the instrument and individual-specific parameters introduces very likely an upward bias in the IV-estimates, both through sorting on the gain and the level. This statement is even more valid if one takes into account that the instrument is negatively related to the quality of kindergarten service provision.

Second, the instrument is likely to be related to test scores through other channels than the number of years spent in kindergarten. It is related through primary school quality, which also leads to an upward bias.

Third, LATE is the treatment effect identified mainly for children from the less privileged families. These children seem to have higher returns than affluent children or the average child. (However, it can happen that affluent children also have such high return, only IV estimation is unable to identify LATE in their case.)[29]

Nevertheless, the effect is found to be significantly positive for disadvantaged children. If the OLS coefficient is indeed downward biased (due to selection into the sample or the sufficiently large covariance between individual-specific abilities and marginal benefit), then the population effects of additional years spent in kindergarten are between 6% and 20% and they are significantly different from zero.

[29] Note that the measurement error in the instrument and the selection into the sample are likely to introduce a downward bias, so the true effects might be even higher or bias in the opposite directions outweigh each other.

Chapter 6: Conclusion

In this thesis I addressed the causal effect of time spent in kindergarten on student achievement. I presented a theoretical model of the choice of time spent in kindergarten. This model and the probability limits of the estimators helped to clarify the channels between test scores, kindergarten attendance and access to kindergarten service. Similar to measuring the causal effect of an additional year spent in school on an individual's lifetime earnings levels, this question is particularly difficult to answer reliably.

The OLS results show that an additional year spent in kindergarten increases student achievement by 6% standard deviation in mathematics and reading (controlling for social, compositional and spatial characteristics of the family and the school-starting age of the child). The IV results show that, keeping the same controls fixed, an additional year spent in kindergarten increases student achievement by 18.4% standard deviation in mathematics and 21.5% standard deviation in reading. If appropriate assumptions are satisfied then the OLS and IV estimates are lower and upper bounds of the population effect, respectively.

The obtained effect is significantly positive, which itself is an important result. The 1st stage estimates by parental education suggest important inequalities across families with different status in access of kindergarten service in Hungary. Those children are more constrained in case of scarcity for whom the highest effect has been estimated. Thus, it is inevitable for educational policy to deal with this inequality problem, especially in the light of current attempts on the extension of compulsory kindergarten attendance.

Let me end with three potential directions for further research.

First, in this thesis I exploited only between-municipality variation. One could make use of within-municipality variation by constructing a consecutive cross-sectional database. In this case bias from non-random distribution of the instrument or its relationship to primary school quality would not be present. Within-municipality variation of teachers across years in very small municipalities could represent a more binding and exogeneous constraint.

Second, by using data for the 4th, 8th and 10th graders one can investigate whether the obtained relationship is stable across grades. Unfortunately test scores obtained from different years cannot be matched to individuals in the HNABC database, therefore one can only analyze different cohorts in the same years to assess the stability across grades. This question would be important from the view of the dynamics in life cycle skill formation.

Third, table14 in the Appendix shows how the results differ if changing the cutoff-value that determines whether a municipality experienced scarcity of kindergarten teachers or not. Higher cutoff-values represent less strict boundaries. There is a pattern of a declining effect of years spent kindergarten and an (in absolute value) increasing effect of the instrument on years spent in kindergarten, if the instrument is stricter. Further research can reconcile this observation with the LATE interpretation discussed in Section 2.3. and the consequences of a gradually weaker instrument (Bound et al., 1995). Further research might find a support for a more appropriate cutoff-value than the 16. percentile.

Appendix

A.Table1. Municipality-level variables for kindergartens ad primary schools, data sources and years

municipality-level variables	Data source	kindergarten	primary school
number of full-time employed teachers	KIR-Stat	1999	2001
> without pedagological qualification	KIR-Stat	1999	2001
> without tertiary qualification	KIR-Stat	1999	2001
> retired	KIR-Stat	1999	2001
> new entrant	KIR-Stat	1999	2001
number of kindergarten places	KIR-Stat	1999	
number of classes	KIR-Stat		
> with below 21 children	KIR-Stat	1999	2001
> with 21-25 children	KIR-Stat	1999	2001
> with 26-30 children	KIR-Stat	1999	2001
> with above 30 children	KIR-Stat	1999	2001
total number of children with age 3, 4 and 5	DEM	1999	
total number of children with age 3, 4 and 6	DEM	1999	
total number of children with age 6-14	DEM		2001
non-wage public expenditures per enrolled child	LGT	1999	2001
wage public expenditures per enrolled child	LGT	1999	2001
non-wage public expenditures per teacher	LGT	1999	
wage public expenditures per teacher	LGT	1999	
non-wage public expenditures per classes	LGT		2001
wage public expenditures per classes	LGT		2001

A.Table 2. Details of sample selection in HNABC2008 for 6th graders, number of observations

all registered students	107654
all registered without special educational needs	100963
> with test scores in reading	95754
> with test scores in mathematics	95757
> with either test scores in reading or mathematics	95790
>> with valid data about kindergarten attendance	83114
>>> with valid data about parental education	79802

A.Table 3. Student-variables and their content

outcome of interest	
Standardized Reading score	with mean 0 and standard deviation 1
Standardized Mathematics score	with mean 0 and standard deviation 1
variable of interest	
Years spent in kindergarten	categories: 0, 1/2, 1, 3/2, 2, 3, 4 (or above) years
control variables	
parental highest educational attainment	
Parental education 0-8th grade (d)	
Parental education vocational (d)	did not obtain upper secondary degree
Parental education secondary (d)	*reference category,* obtained upper secondary degree
Parental education tertiary (d)	finished either university or college
Primary school starting age	
poverty and wealth indicators	
Disadvantaged child (d)	family receives regular childprotection support
Books at home: less than 50 (d)	
Internet at home (d)	
Father employed permanently (d)	
Child works at home permanently (d)	child works permanently on the fields/ in the garden of the family
composition of the household	
Household size	number of individuals living in the household
Lives with biological mother (d)	
Lives with biological father (d)	
Lives with stepmother (d)	
Lives with stepfather (d)	

A.Table3. continues here

settlement's type	
Budapest (d)	*reference category*
County center (d)	
Town (other than county center) (d)	
Village (d)	
settlement's region	
Central Hungary (d)	*reference category*, Budapest and Pest county
Central Transdanubia (d)	Veszprém, Fejér, Komárom-Esztergom counties
Western Transdanubia (d)	Győr-Moson-Sopron, Vas, Zala counties
Southern Transdanubia (d)	Somogy, Tolna, Baranya counties
Northern Hungary (d)	Nógrád, Borsod-Abaúj-Zemplén, Heves counties
Northern Great Plain (d)	Szabolcs-Szatmár-Bereg, Hajdú-Bihar, Jász-Nagykun-Szolnok
Southern Great Plain (d)	Bács-Kiskun, Csongrád, Békés counties

A.Table 4. Descriptive statistics for student-variables

	N	Mean	Stdev.
Standardized reading score	79777	0	1
Standardized mathematics score	79778	0	1
Years spent in kindergarten	79802	3.276	0.751
Years spent in kindergarten <3 (d)	79802	0.105	0.307
Parental education 0-8th grade (d)	79802	0.103	0.304
Parental education vocational (d)	79802	0.303	0.460
Parental education secondary (d)	79802	0.330	0.470
Parental education tertiary (d)	79802	0.264	0.441
Primary school starting age	79802	6.767	0.502
Disadvantaged child (d)	79802	0.251	0.428
Books at home: less than 50 (d)	79802	0.137	0.340
Internet at home (d)	79802	0.684	0.463
Father employed permanently (d)	79802	0.711	0.445
Child works at home permanently (d)	79802	0.131	0.330
Household size	79802	4.412	1.358
Lives with biological mother (d)	79802	0.970	0.170
Lives with biological father (d)	79802	0.795	0.389
Lives with stepmother (d)	79802	0.015	0.121
Lives with stepfather (d)	79802	0.093	0.283
Budapest (d)	79802	0.113	0.317
County center (d)	79802	0.162	0.369
Town (other than county center) (d)	79802	0.351	0.477
Village (d)	79802	0.373	0.484
Central Hungary (d)	79802	0.237	0.425
Central Transdanubia (d)	79802	0.114	0.317
Western Transdanubia (d)	79802	0.101	0.301
Southern Transdanubia (d)	79802	0.096	0.295
Northern Hungary (d)	79802	0.136	0.343
Northern Great Plain (d)	79802	0.183	0.387
Southern Great Plain (d)	79802	0.133	0.340

Notes: Own calculations. Data source: HNABC 2008. Children with special educational needs excluded.(d) denotes dummy variable.

A.Table 5. Average kindergarten attendance by region

6th graders region	**average years spent in kindergarten**	**fraction of those who spent less than 3 years**
Central Hungary	3.285	10.08%
Central Transdanubia	3.346	8.27%
Western Transdanubia	3.334	6.90%
Southern Transdanubia	3.349	7.86%
Northern Hungary	3.158	16.27%
Northern Great Plain	3.194	14.07%
Southern Great Plain	3.339	7.98%

Notes: Own calculations. Data source: HNABC 2008. Children with special educational needs excluded.

A.Table 6. OLS estimates of the effect of time spent in kindergarten on test score; 6th grade, mathematics and reading

	6th grade: mathematics score				6th grade: reading score			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Years spent in kindergarten	0.1641*** [0.005]	0.0674*** [0.004]	0.0607*** [0.004]	0.0613*** [0.004]	0.1763*** [0.005]	0.0762*** [0.004]	0.0667*** [0.004]	0.0594*** [0.004]
Parental education 0-8th grade		-0.9964*** [0.011]	-0.9510*** [0.011]	-0.6343*** [0.013]		-1.0570*** [0.012]	-0.9993*** [0.012]	-0.6417*** [0.013]
Parental education vocational		-0.4707*** [0.008]	-0.4474*** [0.008]	-0.3298*** [0.008]		-0.4649*** [0.008]	-0.4374*** [0.008]	-0.3074*** [0.008]
Parental education tertiary		0.4205*** [0.008]	0.3920*** [0.008]	0.3525*** [0.008]		0.3988*** [0.008]	0.3667*** [0.008]	0.3283*** [0.008]
Settlement type: County center			0.0590*** [0.015]	0.0406*** [0.015]			0.0418*** [0.015]	0.0235 [0.014]
Settlement type: Town			-0.0850*** [0.013]	-0.0829*** [0.013]			-0.0854*** [0.012]	-0.0773*** [0.012]
Settlement type: Village			-0.1205*** [0.014]	-0.0957*** [0.013]			-0.1570*** [0.013]	-0.1194*** [0.013]
Region: Central Transdanubia			-0.0346*** [0.013]	-0.0366*** [0.012]			-0.0074 [0.012]	-0.0111 [0.012]
Region: Western Transdanubia			0.0187 [0.013]	0.0145 [0.013]			0.0370*** [0.013]	0.0325*** [0.012]
Region: Southern Transdanubia			-0.0480*** [0.013]	-0.0123 [0.013]			-0.0302** [0.013]	0.0062 [0.013]
Region: Northern Hungary			-0.1315*** [0.012]	-0.0832*** [0.012]			-0.1600*** [0.012]	-0.1076*** [0.012]
Region: Northern Great Plain			-0.0875*** [0.012]	-0.0221* [0.012]			-0.1150*** [0.011]	-0.0434*** [0.011]
Region: Southern Great Plain			-0.0593*** [0.012]	-0.0284** [0.012]			-0.0429*** [0.012]	-0.0113 [0.012]
Father employed permanently				0.0981*** [0.010]				0.0995*** [0.010]
Primary school starting age				-0.1272*** [0.006]				-0.1121*** [0.006]
Disadvantaged child				-0.1079*** [0.008]				-0.1063*** [0.008]
Household size				-0.0185*** [0.002]				-0.0373*** [0.003]
Lives with biological mother				0.0562*** [0.021]				0.0735*** [0.022]
Lives with biological father				-0.0134 [0.009]				-0.0184** [0.009]
Lives with stepmother				-0.1759*** [0.029]				-0.1244*** [0.031]
Lives with stepfather				-0.0931*** [0.012]				-0.0549*** [0.012]
Books at home: less than 50				-0.2379*** [0.010]				-0.2941*** [0.010]
Internet at home				0.1463*** [0.008]				0.1447*** [0.008]
Child works at home permanently				-0.1446*** [0.009]				-0.1763*** [0.009]
Constant	-0.4934*** [0.016]	-0.0418*** [0.015]	0.0897*** [0.018]	0.8032*** [0.049]	-0.5294*** [0.017]	-0.0565*** [0.015]	0.0990*** [0.017]	0.7959*** [0.049]
dummies for imputed missings	yes	yes	yes	yes	yes	yes	yes	yes
Observations	79,778	79,778	79,778	79,778	79,777	79,777	79,777	79,777
R-squared	0.016	0.215	0.222	0.249	0.018	0.224	0.234	0.267

*Notes: Own calculations. Data source: HNABC 2008, KIR-Stat 1996-2001, DEM 1996-2001. Children with special needs excluded. Reference categories: Parental education: secondary, Settlement type: Budapest, Region: Central-Hungary. Scores are measured in standard deviations. Robust standard errors in parentheses. Significance levels: *** 1%, ** 5%, * 10%. Missing variables imputed with the sample average and control variables are included of dummy-variables that denote missing imputed observations with 1.*

A.Table 7. IV estimates of the effect of time spent in kindergarten on test score; 6th grade, mathematics and reading

	2008 6th grade: mathematics score				2008 6th grade: reading score			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Years spent in kindergarten	1.1745***	0.5283***	0.3286***	0.1844***	1.2786***	0.6472***	0.3785***	0.2153***
	[0.041]	[0.038]	[0.043]	[0.044]	[0.043]	[0.040]	[0.044]	[0.044]
Parental education 0-8th grade		-0.7776***	-0.8326***	-0.6019***		-0.7846***	-0.8599***	-0.5992***
		[0.022]	[0.022]	[0.018]		[0.023]	[0.023]	[0.018]
Parental education vocational		-0.4176***	-0.4186***	-0.3235***		-0.3982***	-0.4030***	-0.2985***
		[0.009]	[0.009]	[0.008]		[0.010]	[0.009]	[0.008]
Parental education tertiary		0.4190***	0.3941***	0.3532***		0.3971***	0.3694***	0.3295***
		[0.009]	[0.008]	[0.008]		[0.009]	[0.008]	[0.008]
Settlement type: County center			0.0553***	0.0422***			0.0363**	0.0239
			[0.016]	[0.015]			[0.015]	[0.015]
Settlement type: Town			-0.0751***	-0.0800***			-0.0748***	-0.0751***
			[0.013]	[0.013]			[0.013]	[0.012]
Settlement type: Village			-0.0948***	-0.0904***			-0.1283***	-0.1145***
			[0.014]	[0.014]			[0.014]	[0.013]
Region: Central Transdanubia			-0.0674***	-0.0535***			-0.0424***	-0.0289**
			[0.014]	[0.013]			[0.014]	[0.013]
Region: Western Transdanubia			-0.0052	0.0028			0.0103	0.0192
			[0.014]	[0.013]			[0.014]	[0.013]
Region: Southern Transdanubia			-0.0833***	-0.0324**			-0.0705***	-0.0183
			[0.015]	[0.015]			[0.015]	[0.014]
Region: Northern Hungary			-0.1212***	-0.0848***			-0.1472***	-0.1084***
			[0.013]	[0.012]			[0.013]	[0.012]
Region: Northern Great Plain			-0.0853***	-0.0290**			-0.1112***	-0.0508***
			[0.012]	[0.012]			[0.012]	[0.012]
Region: Southern Great Plain			-0.0863***	-0.0456***			-0.0733***	-0.0317**
			[0.013]	[0.013]			[0.013]	[0.013]
Father employed permanently				0.0880***				0.0862***
				[0.010]				[0.010]
Primary school starting age				-0.1749***				-0.1732***
				[0.018]				[0.018]
Disadvantaged child				-0.0974***				-0.0931***
				[0.009]				[0.009]
Household size				-0.0124***				-0.0295***
				[0.003]				[0.003]
Lives with biological mother				0.0488**				0.0641***
				[0.021]				[0.022]
Lives with biological father				-0.0123				-0.0165*
				[0.009]				[0.009]
Lives with stepmother				-0.1690***				-0.1162***
				[0.030]				[0.031]
Lives with stepfather				-0.1005***				-0.0646***
				[0.013]				[0.013]
Books at home: less than 50				-0.2228***				-0.2755***
				[0.012]				[0.012]
Internet at home				0.1360***				0.1314***
				[0.009]				[0.009]
Child works at home permanently				-0.1401***				-0.1697***
				[0.010]				[0.010]
Constant	-3.8029***	-1.5900***	-0.8107***	0.7121***	-4.1402***	-1.9750***	-0.9495***	0.6865***
	[0.134]	[0.129]	[0.145]	[0.058]	[0.140]	[0.135]	[0.148]	[0.059]
dummies for imputed missings	yes	yes	yes	yes	yes	yes	yes	yes
Observations	79,507	79,507	79,507	79,507	79,506	79,506	79,506	79,506
R-squared		0.096	0.182	0.242		0.038	0.179	0.254

*Notes: Own calculations. Data source: HNABC 2008, KIR-Stat 1996-2001, DEM 1996-2001. Children with special needs excluded. Reference categories: Parental education: secondary, Settlement type: Budapest, Region: Central-Hungary. Scores are measured in standard deviations. Robust standard errors in parentheses. Significance levels: *** 1%, ** 5%, * 10%. Instrument: kindergarten teacher over kindergarten-aged (3-5 aged) population ratio is in the lowest sextile in the municipality the child lived at age 4. Missing variables imputed with the sample average and control variables are included of dummy-variables that denote missing imputed observations with 1.*

A.Table 8. IV 1st stage estimates on years spent in kindergarten, 6th grade

	2008 6th grade: Years spent in kindergarten			
	(1)	(2)	(3)	(4)
IV: Teachers / children aged 3-5 < 16.percentile	-0.3023*** [0.008]	-0.2437*** [0.008]	-0.2186*** [0.008]	-0.2084*** [0.008]
Parental education 0-8th grade		-0.4317*** [0.011]	-0.4206*** [0.011]	-0.2622*** [0.013]
Parental education vocational		-0.0973*** [0.006]	-0.0985*** [0.006]	-0.0484*** [0.006]
Parental education tertiary		-0.0091 [0.006]	-0.0101 [0.006]	-0.0062 [0.006]
Settlement type: County center			0.0382*** [0.013]	0.0264** [0.012]
Settlement type: Town			0.0061 [0.011]	0.0301*** [0.011]
Settlement type: Village			-0.0070 [0.012]	0.0532*** [0.011]
Region: Central Transdanubia			0.0766*** [0.011]	0.0700*** [0.010]
Region: Western Transdanubia			0.0515*** [0.011]	0.0446*** [0.011]
Region: Southern Transdanubia			0.0996*** [0.011]	0.1181*** [0.011]
Region: Northern Hungary			-0.0553*** [0.011]	-0.0245** [0.011]
Region: Northern Great Plain			-0.0146 [0.010]	0.0329*** [0.010]
Region: Southern Great Plain			0.0705*** [0.010]	0.0959*** [0.010]
Father employed permanently				0.0733*** [0.009]
Primary school starting age				0.3911*** [0.005]
Disadvantaged child				-0.0741*** [0.007]
Household size				-0.0482*** [0.002]
Lives with biological mother				0.0746*** [0.019]
Lives with biological father				-0.0109 [0.008]
Lives with stepmother				-0.0300 [0.028]
Lives with stepfather				0.0595*** [0.010]
Books at home: less than 50				-0.1170*** [0.010]
Internet at home				0.0750*** [0.006]
Child works at home permanently				-0.0409*** [0.008]
Constant	3.3267*** [0.003]	3.3933*** [0.004]	3.3600*** [0.008]	0.7042*** [0.041]
dummies for imputed missings	yes	yes	yes	yes
Observations	79,531	79,531	79,531	79,531
R-squared	0.023	0.051	0.055	0.140

*Notes: Own calculations. Data source: HNABC 2008, KIR-Stat 1996-2001, DEM 1996-2001. Children with special needs excluded. Reference categories: Parental education: secondary, Settlement type: Budapest, Region: Central-Hungary. Scores are measured in standard deviations. Robust standard errors in parentheses. Significance levels: *** 1%, ** 5%, * 10%. Instrument: kindergarten teacher over kindergarten-aged (3-5 aged) population ratio is in the lowest sextile in the municipality the child lived at age 4. Missing variables imputed with the sample average and control variables are included of dummy-variables that denote missing imputed observations with 1.*

A.Table 9. IV estimates of the effect of time spent in kindergarten on test score; 6th grade, mathematics and reading, instrument appropriate to different robustness checks

	#1: places / 3-5		#2: teachers / 3-6		#2: places / 3-6	
	mathematics	reading	mathematics	reading	mathematics	reading
Years spent in kindergarten	0.2642***	0.2241***	0.1609***	0.2001***	0.2321***	0.1946***
	[0.036]	[0.036]	[0.043]	[0.043]	[0.036]	[0.036]
Parental education 0-8th grade	-0.5800***	-0.5966***	-0.6083***	-0.6033***	-0.5887***	-0.6047***
	[0.017]	[0.017]	[0.018]	[0.018]	[0.016]	[0.017]
Parental education vocational	-0.3193***	-0.2980***	-0.3247***	-0.2993***	-0.3210***	-0.2996***
	[0.008]	[0.008]	[0.008]	[0.008]	[0.008]	[0.008]
Parental education tertiary	0.3536***	0.3296***	0.3531***	0.3294***	0.3534***	0.3294***
	[0.008]	[0.008]	[0.008]	[0.008]	[0.008]	[0.008]
Settlement type: County center	0.0410***	0.0239	0.0426***	0.0242*	0.0415***	0.0243*
	[0.015]	[0.015]	[0.015]	[0.015]	[0.015]	[0.015]
Settlement type: Town	-0.0800***	-0.0751***	-0.0800***	-0.0751***	-0.0800***	-0.0751***
	[0.013]	[0.012]	[0.013]	[0.012]	[0.013]	[0.012]
Settlement type: Village	-0.0890***	-0.1144***	-0.0909***	-0.1148***	-0.0896***	-0.1150***
	[0.014]	[0.013]	[0.014]	[0.013]	[0.014]	[0.013]
Region: Central Transdanubia	-0.0608***	-0.0297**	-0.0514***	-0.0275**	-0.0578***	-0.0269**
	[0.013]	[0.013]	[0.013]	[0.013]	[0.013]	[0.013]
Region: Western Transdanubia	-0.0031	0.0180	0.0045	0.0203	-0.0008	0.0201
	[0.013]	[0.013]	[0.013]	[0.013]	[0.013]	[0.013]
Region: Southern Transdanubia	-0.0431***	-0.0194	-0.0292**	-0.0161	-0.0388***	-0.0152
	[0.014]	[0.014]	[0.014]	[0.014]	[0.014]	[0.014]
Region: Northern Hungary	-0.0830***	-0.1082***	-0.0854***	-0.1088***	-0.0838***	-0.1089***
	[0.013]	[0.012]	[0.012]	[0.012]	[0.012]	[0.012]
Region: Northern Great Plain	-0.0314***	-0.0511***	-0.0283**	-0.0503***	-0.0304**	-0.0501***
	[0.012]	[0.012]	[0.012]	[0.012]	[0.012]	[0.012]
Region: Southern Great Plain	-0.0548***	-0.0328***	-0.0429***	-0.0300**	-0.0511***	-0.0294**
	[0.013]	[0.013]	[0.013]	[0.013]	[0.013]	[0.013]
Father employed permanently	0.0821***	0.0858***	0.0898***	0.0873***	0.0845***	0.0880***
	[0.010]	[0.010]	[0.010]	[0.010]	[0.010]	[0.010]
Primary school starting age	-0.2059***	-0.1767***	-0.1657***	-0.1673***	-0.1935***	-0.1652***
	[0.015]	[0.015]	[0.018]	[0.018]	[0.015]	[0.015]
Disadvantaged child	-0.0909***	-0.0922***	-0.0995***	-0.0944***	-0.0936***	-0.0947***
	[0.009]	[0.009]	[0.009]	[0.009]	[0.009]	[0.009]
Household size	-0.0084***	-0.0291***	-0.0135***	-0.0303***	-0.0100***	-0.0305***
	[0.003]	[0.003]	[0.003]	[0.003]	[0.003]	[0.003]
Lives with biological mother	0.0430**	0.0633***	0.0505**	0.0652***	0.0453**	0.0654***
	[0.021]	[0.022]	[0.021]	[0.022]	[0.021]	[0.022]
Lives with biological father	-0.0111	-0.0161*	-0.0127	-0.0167*	-0.0116	-0.0165*
	[0.009]	[0.009]	[0.009]	[0.009]	[0.009]	[0.009]
Lives with stepmother	-0.1665***	-0.1159***	-0.1698***	-0.1167***	-0.1675***	-0.1169***
	[0.030]	[0.031]	[0.030]	[0.031]	[0.030]	[0.031]
Lives with stepfather	-0.1052***	-0.0650***	-0.0992***	-0.0636***	-0.1034***	-0.0631***
	[0.013]	[0.013]	[0.013]	[0.013]	[0.013]	[0.013]
Books at home: less than 50	-0.2132***	-0.2743***	-0.2255***	-0.2773***	-0.2170***	-0.2779***
	[0.011]	[0.011]	[0.012]	[0.012]	[0.011]	[0.011]
Internet at home	0.1296***	0.1309***	0.1379***	0.1327***	0.1323***	0.1334***
	[0.008]	[0.008]	[0.008]	[0.009]	[0.008]	[0.008]
Child works at home permanently	-0.1363***	-0.1695***	-0.1411***	-0.1704***	-0.1377***	-0.1708***
	[0.010]	[0.010]	[0.010]	[0.010]	[0.010]	[0.010]
Constant	0.6545***	0.6799***	0.7291***	0.6975***	0.6778***	0.7012***
	[0.056]	[0.056]	[0.058]	[0.058]	[0.056]	[0.055]
dummies for imputed missings	yes	yes	yes	yes	yes	yes
Observations	79,501	79,500	79,509	79,508	79,503	79,502
R-squared	0.228	0.253	0.244	0.256	0.234	0.257

*Notes: Own calculations. Data source: HNABC 2008, KIR-Stat 1996-2001, DEM 1996-2001. Children with special needs excluded. Reference categories: Parental education: secondary, Settlement type: Budapest, Region: Central-Hungary. Scores are measured in standard deviations. Robust standard errors in parentheses. Significance levels: *** 1%, ** 5%, * 10%. Instruments: Robustness Check #1: kindergarten places over kindergarten-aged (3-5 aged) population ratio is in the lowest sextile in the municipality the child lived at age 4.Robustness Check #2: kindergarten teachers over kindergarten-aged (3-6 aged) population ratio is in the lowest sextile in the municipality the child lived at age 4.Robustness Check #3: kindergarten places over kindergarten-aged (3-6 aged) population ratio is in the lowest sextile in the municipality the child lived at age 4. Missing variables imputed with the sample average and control variables are included of dummy-variables that denote missing imputed observations with 1.*

A.Table 10. IV 1st stage estimates on years spent in kindergarten, 6th grade, instrument appropriate to different robustness checks

	#1: places / 3-5	#2: teachers / 3-6	#2: places / 3-6
appropriate IV	-0.2503*** [0.008]	-0.2114*** [0.008]	-0.2512*** [0.008]
Parental education 0-8th grade	-0.2629*** [0.013]	-0.2616*** [0.013]	-0.2635*** [0.013]
Parental education vocational	-0.0500*** [0.006]	-0.0483*** [0.006]	-0.0497*** [0.006]
Parental education tertiary	-0.0037 [0.006]	-0.0061 [0.006]	-0.0034 [0.006]
Settlement type: County center	0.0538*** [0.012]	0.0194 [0.012]	0.0545*** [0.012]
Settlement type: Town	0.0709*** [0.011]	0.0251** [0.011]	0.0710*** [0.011]
Settlement type: Village	0.0782*** [0.012]	0.0468*** [0.011]	0.0796*** [0.012]
Region: Central Transdanubia	0.0371*** [0.011]	0.0773*** [0.010]	0.0361*** [0.011]
Region: Western Transdanubia	0.0153 [0.011]	0.0501*** [0.010]	0.0147 [0.011]
Region: Southern Transdanubia	0.0899*** [0.011]	0.1234*** [0.011]	0.0890*** [0.011]
Region: Northern Hungary	-0.0448*** [0.011]	-0.0151 [0.011]	-0.0448*** [0.011]
Region: Northern Great Plain	0.0139 [0.010]	0.0393*** [0.010]	0.0126 [0.010]
Region: Southern Great Plain	0.0592*** [0.010]	0.1035*** [0.010]	0.0596*** [0.010]
Father employed permanently	0.0738*** [0.009]	0.0739*** [0.009]	0.0744*** [0.009]
Primary school starting age	0.3893*** [0.005]	0.3913*** [0.005]	0.3893*** [0.005]
Disadvantaged child	-0.0736*** [0.007]	-0.0743*** [0.007]	-0.0737*** [0.007]
Household size	-0.0476*** [0.002]	-0.0480*** [0.002]	-0.0478*** [0.002]
Lives with biological mother	0.0741*** [0.019]	0.0744*** [0.019]	0.0749*** [0.019]
Lives with biological father	-0.0112 [0.008]	-0.0108 [0.008]	-0.0110 [0.008]
Lives with stepmother	-0.0300 [0.028]	-0.0321 [0.028]	-0.0311 [0.028]
Lives with stepfather	0.0567*** [0.010]	0.0599*** [0.010]	0.0568*** [0.010]
Books at home: less than 50	-0.1161*** [0.009]	-0.1167*** [0.010]	-0.1156*** [0.009]
Internet at home	0.0749*** [0.006]	0.0740*** [0.006]	0.0748*** [0.006]
Child works at home permanently	-0.0416*** [0.008]	-0.0407*** [0.008]	-0.0421*** [0.008]
Constant	0.7138*** [0.040]	0.7031*** [0.041]	0.7132*** [0.040]
dummies for imputed missings	yes	yes	yes
Observations	79,525	79,533	79,527
R-squared	0.144	0.140	0.144

*Notes: Own calculations. Data source: HNABC 2008, KIR-Stat 1996-2001, DEM 1996-2001. Children with special needs excluded. Reference categories: Parental education: secondary, Settlement type: Budapest, Region: Central-Hungary. Scores are measured in standard deviations. Robust standard errors in parentheses. Significance levels: *** 1%, ** 5%, * 10%. Instruments: Robustness Check #1: kindergarten places over kindergarten-aged (3-5 aged) population ratio is in the lowest sextile in the municipality the child lived at age 4.Robustness Check #2: kindergarten teachers over kindergarten-aged (3-6 aged) population ratio is in the lowest sextile in the municipality the child lived at age 4.Robustness Check #3: kindergarten places over kindergarten-aged (3-6 aged) population ratio is in the lowest sextile in the municipality the child lived at age 4. Missing variables imputed with the sample average and control variables are included of dummy-variables that denote missing imputed observations with 1.*

A.Table 11. OLS estimates of the effect of time spent in kindergarten on test score; 6th grade, mathematics and reading, by parental education

	mathematics				reading			
parental education:	grade 0-8 (1)	vocational (2)	secondary (3)	tertiary (4)	grade 0-8 (5)	vocational (6)	secondary (7)	tertiary (8)
Years spent in kindergarten	0.0591***	0.0518***	0.0655***	0.0587***	0.0775***	0.0523***	0.0525***	0.0444***
	[0.011]	[0.008]	[0.008]	[0.010]	[0.011]	[0.008]	[0.008]	[0.009]
Settlement type: County center	0.0909	0.0354	0.0469*	0.0790***	0.0390	-0.0251	0.0244	0.0741***
	[0.067]	[0.033]	[0.026]	[0.025]	[0.068]	[0.033]	[0.025]	[0.024]
Settlement type: Town	-0.0217	-0.0932***	-0.0399*	-0.0894***	-0.0565	-0.1200***	-0.0531**	-0.0619***
	[0.057]	[0.029]	[0.021]	[0.021]	[0.059]	[0.029]	[0.021]	[0.020]
Settlement type: Village	-0.0080	-0.0864***	-0.0526**	-0.1561***	-0.0751	-0.1465***	-0.0973***	-0.1423***
	[0.057]	[0.029]	[0.023]	[0.023]	[0.059]	[0.029]	[0.022]	[0.022]
Region: Central Transdanubia	-0.0222	-0.0012	-0.0635***	-0.0386	0.0357	0.0127	-0.0230	-0.0267
	[0.044]	[0.023]	[0.021]	[0.024]	[0.046]	[0.023]	[0.021]	[0.023]
Region: Western Transdanubia	0.0147	0.0416*	0.0456**	-0.0511**	0.0688	0.0800***	0.0488**	-0.0440*
	[0.049]	[0.024]	[0.021]	[0.025]	[0.050]	[0.023]	[0.021]	[0.023]
Region: Southern Transdanubia	-0.0136	0.0092	0.0240	-0.0938***	0.0542	0.0250	0.0133	-0.0575**
	[0.043]	[0.023]	[0.023]	[0.027]	[0.044]	[0.024]	[0.023]	[0.025]
Region: Northern Hungary	-0.2031***	-0.0462**	-0.0279	-0.1253***	-0.2874***	-0.0830***	-0.0416**	-0.0972***
	[0.040]	[0.022]	[0.021]	[0.024]	[0.041]	[0.023]	[0.021]	[0.023]
Region: Northern Great Plain	-0.0967**	0.0442**	0.0019	-0.1048***	-0.1202***	0.0253	-0.0317	-0.1063***
	[0.038]	[0.021]	[0.020]	[0.023]	[0.039]	[0.021]	[0.020]	[0.022]
Region: Southern Great Plain	0.0503	0.0269	-0.0219	-0.1378***	0.0983**	0.0392*	-0.0251	-0.0931***
	[0.042]	[0.022]	[0.020]	[0.023]	[0.043]	[0.022]	[0.020]	[0.022]
Father employed permanently	0.0583***	0.0885***	0.0758***	0.1725***	0.0820***	0.0833***	0.0751***	0.1517***
	[0.021]	[0.015]	[0.019]	[0.025]	[0.022]	[0.016]	[0.018]	[0.024]
Primary school starting age	-0.1480***	-0.1357***	-0.1095***	-0.1211***	-0.1235***	-0.1167***	-0.0958***	-0.1054***
	[0.018]	[0.011]	[0.011]	[0.013]	[0.019]	[0.011]	[0.011]	[0.012]
Disadvantaged child	-0.1012***	-0.0982***	-0.0882***	-0.2128***	-0.0731***	-0.0914***	-0.1058***	-0.1888***
	[0.023]	[0.013]	[0.016]	[0.025]	[0.023]	[0.013]	[0.015]	[0.024]
Household size	-0.0454***	-0.0318***	-0.0100**	0.0164***	-0.0619***	-0.0512***	-0.0264***	-0.0023
	[0.005]	[0.004]	[0.005]	[0.006]	[0.006]	[0.004]	[0.005]	[0.005]
Lives with biological mother	-0.0335	0.0124	0.1194***	0.1428***	0.0040	0.0185	0.0969**	0.2160***
	[0.048]	[0.033]	[0.043]	[0.051]	[0.052]	[0.035]	[0.045]	[0.050]
Lives with biological father	-0.0843***	-0.0663***	-0.0327**	0.0521***	-0.0640**	-0.0613***	-0.0362**	0.0272
	[0.029]	[0.016]	[0.016]	[0.019]	[0.029]	[0.017]	[0.016]	[0.018]
Lives with stepmother	-0.0819	-0.1459***	-0.1904***	-0.2399***	-0.0309	-0.1364***	-0.1477**	-0.1114
	[0.074]	[0.048]	[0.055]	[0.071]	[0.077]	[0.052]	[0.057]	[0.072]
Lives with stepfather	-0.0065	-0.0257	-0.1453***	-0.1913***	0.0519	0.0067	-0.1069***	-0.1530***
	[0.040]	[0.021]	[0.021]	[0.027]	[0.040]	[0.021]	[0.021]	[0.026]
Books at home: less than 50	-0.1104***	-0.2598***	-0.3081***	-0.4943***	-0.1683***	-0.3064***	-0.3713***	-0.5859***
	[0.020]	[0.014]	[0.025]	[0.058]	[0.021]	[0.014]	[0.024]	[0.053]
Internet at home	0.1655***	0.1392***	0.1397***	0.1899***	0.1265***	0.1544***	0.1409***	0.1457***
	[0.024]	[0.012]	[0.014]	[0.021]	[0.025]	[0.012]	[0.013]	[0.021]
Child works at home permanently	-0.1237***	-0.1160***	-0.1653***	-0.1719***	-0.1511***	-0.1662***	-0.1720***	-0.2120***
	[0.025]	[0.015]	[0.017]	[0.021]	[0.026]	[0.016]	[0.017]	[0.020]
Constant	0.5003***	0.6684***	0.5726***	0.8018***	0.3267**	0.6879***	0.6616***	0.7978***
	[0.148]	[0.087]	[0.088]	[0.100]	[0.151]	[0.088]	[0.089]	[0.096]
dummies for imputed missings	yes	yes	yes	yes	yes	yes	yes	yes
Observations	8,226	24,210	26,310	21,032	8,224	24,208	26,310	21,035
R-squared	0.077	0.062	0.035	0.053	0.108	0.081	0.043	0.051

*Notes: Own calculations. Data source: HNABC 2008, KIR-Stat 1996-2001, DEM 1996-2001. Children with special needs excluded. Reference categories: Parental education: secondary, Settlement type: Budapest, Region: Central-Hungary. Scores are measured in standard deviations. Robust standard errors in parentheses. Significance levels: *** 1%, ** 5%, * 10%. Missing variables imputed with the sample average and control variables are included of dummy-variables that denote missing imputed observations with 1.*

A.Table12. IV estimates of the effect of time spent in kindergarten on test score; 6th grade, mathematics and reading, by parental education

	mathematics				reading			
parental education:	grade 0-8 (1)	vocational (2)	secondary (3)	tertiary (4)	grade 0-8 (5)	vocational (6)	secondary (7)	tertiary (8)
Years spent in kindergarten	0.139** [0.066]	0.217*** [0.061]	0.119 [0.117]	0.198 [0.366]	0.143** [0.069]	0.251*** [0.062]	0.078 [0.116]	0.439 [0.366]
Settlement type: County center	0.095 [0.068]	0.045 [0.034]	0.049* [0.026]	0.078*** [0.028]	0.045 [0.069]	-0.015 [0.033]	0.025 [0.025]	0.064** [0.027]
Settlement type: Town	-0.005 [0.059]	-0.081*** [0.029]	-0.038* [0.021]	-0.092*** [0.025]	-0.041 [0.060]	-0.107*** [0.029]	-0.052** [0.021]	-0.075*** [0.025]
Settlement type: Village	0.005 [0.058]	-0.071** [0.030]	-0.050** [0.023]	-0.155*** [0.024]	-0.062 [0.060]	-0.129*** [0.030]	-0.097*** [0.022]	-0.144*** [0.023]
Region: Central Transdanubia	-0.056 [0.050]	-0.031 [0.024]	-0.069*** [0.022]	-0.049 [0.031]	0.006 [0.051]	-0.018 [0.024]	-0.024 [0.021]	-0.050 [0.031]
Region: Western Transdanubia	-0.017 [0.054]	0.023 [0.024]	0.042* [0.021]	-0.060** [0.030]	0.041 [0.055]	0.058** [0.024]	0.048** [0.021]	-0.064** [0.029]
Region: Southern Transdanubia	-0.041 [0.049]	-0.022 [0.025]	0.018 [0.025]	-0.109*** [0.042]	0.027 [0.050]	-0.011 [0.026]	0.010 [0.025]	-0.095** [0.042]
Region: Northern Hungary	-0.193*** [0.042]	-0.044* [0.023]	-0.032 [0.021]	-0.142*** [0.038]	-0.281*** [0.043]	-0.080*** [0.024]	-0.046** [0.021]	-0.133*** [0.038]
Region: Northern Great Plain	-0.099*** [0.038]	0.035 [0.021]	-0.001 [0.020]	-0.122*** [0.044]	-0.125*** [0.039]	0.016 [0.022]	-0.033* [0.020]	-0.148*** [0.043]
Region: Southern Great Plain	0.031 [0.045]	-0.001 [0.023]	-0.028 [0.023]	-0.154*** [0.042]	0.080* [0.046]	0.007 [0.024]	-0.028 [0.023]	-0.133*** [0.041]
Father employed permanently	0.051** [0.023]	0.076*** [0.016]	0.074*** [0.019]	0.159*** [0.037]	0.076*** [0.024]	0.066*** [0.016]	0.075*** [0.019]	0.119*** [0.037]
Primary school starting age	-0.166*** [0.023]	-0.197*** [0.025]	-0.132** [0.052]	-0.181 [0.157]	-0.139*** [0.025]	-0.192*** [0.026]	-0.107** [0.051]	-0.276* [0.157]
Disadvantaged child	-0.094*** [0.023]	-0.088*** [0.013]	-0.084*** [0.017]	-0.196*** [0.056]	-0.065*** [0.024]	-0.079*** [0.014]	-0.103*** [0.017]	-0.135** [0.056]
Household size	-0.042*** [0.006]	-0.024*** [0.005]	-0.008 [0.007]	0.024 [0.020]	-0.058*** [0.006]	-0.041*** [0.005]	-0.026*** [0.007]	0.019 [0.020]
Lives with biological mother	-0.045 [0.048]	0.008 [0.033]	0.112** [0.044]	0.134** [0.068]	0.002 [0.052]	0.014 [0.036]	0.093** [0.046]	0.167** [0.069]
Lives with biological father	-0.083*** [0.029]	-0.064*** [0.017]	-0.032** [0.016]	0.051** [0.020]	-0.065** [0.029]	-0.056*** [0.017]	-0.036** [0.016]	0.023 [0.019]
Lives with stepmother	-0.084 [0.074]	-0.125** [0.049]	-0.194*** [0.055]	-0.230*** [0.072]	-0.024 [0.078]	-0.109** [0.053]	-0.150*** [0.058]	-0.125* [0.076]
Lives with stepfather	-0.014 [0.041]	-0.037* [0.021]	-0.148*** [0.022]	-0.197*** [0.036]	0.040 [0.041]	-0.005 [0.022]	-0.108*** [0.022]	-0.176*** [0.035]
Books at home: less than 50	-0.104*** [0.021]	-0.239*** [0.016]	-0.302*** [0.027]	-0.468*** [0.083]	-0.165*** [0.022]	-0.282*** [0.016]	-0.369*** [0.026]	-0.515*** [0.082]
Internet at home	0.157*** [0.025]	0.124*** [0.013]	0.137*** [0.016]	0.182*** [0.028]	0.119*** [0.026]	0.136*** [0.013]	0.139*** [0.016]	0.127*** [0.028]
Child works at home permanently	-0.125*** [0.025]	-0.111*** [0.016]	-0.163*** [0.018]	-0.161*** [0.032]	-0.156*** [0.026]	-0.158*** [0.016]	-0.170*** [0.018]	-0.185*** [0.032]
Constant	0.377** [0.182]	0.530*** [0.101]	0.545*** [0.104]	0.741*** [0.177]	0.224 [0.183]	0.524*** [0.103]	0.654*** [0.104]	0.647*** [0.179]
dummies for imputed missings	yes	yes	yes	yes	yes	yes	yes	yes
Observations	8,192	24,102	26,221	20,992	8,190	24,100	26,221	20,995
R-squared	0.071	0.043	0.034	0.043	0.103	0.054	0.043	

*Notes: Own calculations. Data source: HNABC 2008, KIR-Stat 1996-2001, DEM 1996-2001. Children with special needs excluded. Reference categories: Parental education: secondary, Settlement type: Budapest, Region: Central-Hungary. Scores are measured in standard deviations. Robust standard errors in parentheses. Significance levels: *** 1%, ** 5%, * 10%. Instrument: kindergarten teacher over kindergarten-aged (3-5 aged) population ratio is in the lowest sextile in the municipality the child lived at age 4. Missing variables imputed with the sample average and control variables are included of dummy-variables that denote missing imputed observations with 1.*

A.Table 13. IV 1st stage estimates on years spent in kindergarten, 6th grade, by parental education

parental education:	2008 6th grade: Years spent in kindergarten grade 0-8 (1)	vocational (2)	secondary (3)	tertiary (4)
IV: Teachers / children aged 3-5 < 16.percentile	-0.347*** [0.024]	-0.244*** [0.013]	-0.138*** [0.014]	-0.063*** [0.018]
Settlement type: County center	-0.003 [0.071]	-0.011 [0.029]	0.006 [0.020]	0.036* [0.019]
Settlement type: Town	-0.111* [0.063]	-0.001 [0.026]	0.025 [0.017]	0.045*** [0.017]
Settlement type: Village	0.005 [0.063]	0.031 [0.026]	0.023 [0.018]	0.028 [0.019]
Region: Central Transdanubia	0.277*** [0.048]	0.092*** [0.020]	0.027 [0.017]	0.048*** [0.019]
Region: Western Transdanubia	0.261*** [0.051]	0.054*** [0.021]	0.017 [0.016]	0.040** [0.019]
Region: Southern Transdanubia	0.296*** [0.048]	0.133*** [0.020]	0.074*** [0.017]	0.084*** [0.019]
Region: Northern Hungary	-0.131*** [0.046]	-0.039* [0.021]	0.006 [0.016]	0.076*** [0.019]
Region: Northern Great Plain	0.026 [0.044]	0.029 [0.019]	0.017 [0.016]	0.100*** [0.017]
Region: Southern Great Plain	0.180*** [0.048]	0.113*** [0.019]	0.070*** [0.016]	0.090*** [0.017]
Father employed permanently	0.118*** [0.023]	0.065*** [0.014]	0.011 [0.015]	0.072*** [0.020]
Primary school starting age	0.230*** [0.019]	0.378*** [0.009]	0.431*** [0.008]	0.429*** [0.009]
Disadvantaged child	-0.059** [0.024]	-0.056*** [0.011]	-0.057*** [0.013]	-0.138*** [0.021]
Household size	-0.042*** [0.006]	-0.049*** [0.004]	-0.039*** [0.004]	-0.053*** [0.005]
Lives with biological mother	0.120** [0.053]	0.045 [0.029]	0.082** [0.037]	0.125*** [0.045]
Lives with biological father	-0.053* [0.029]	-0.023* [0.014]	0.004 [0.012]	0.014 [0.014]
Lives with stepmother	0.109 [0.076]	-0.103** [0.045]	-0.042 [0.049]	0.036 [0.069]
Lives with stepfather	0.086** [0.039]	0.056*** [0.018]	0.045*** [0.016]	0.065*** [0.020]
Books at home: less than 50	-0.082*** [0.021]	-0.120*** [0.013]	-0.091*** [0.020]	-0.161*** [0.052]
Internet at home	0.100*** [0.024]	0.078*** [0.010]	0.071*** [0.011]	0.049*** [0.016]
Child works at home permanently	0.021 [0.026]	-0.042*** [0.013]	-0.056*** [0.014]	-0.066*** [0.017]
Constant	1.517*** [0.153]	0.813*** [0.072]	0.457*** [0.068]	0.396*** [0.078]
dummies for imputed missings	yes	yes	yes	yes
Observations	8,196	24,106	26,227	21,002
R-squared	0.116	0.116	0.115	0.122

*Notes: Own calculations. Data source: HNABC 2008, KIR-Stat 1996-2001, DEM 1996-2001. Children with special needs excluded. Reference categories: Parental education: secondary, Settlement type: Budapest, Region: Central-Hungary. Scores are measured in standard deviations. Robust standard errors in parentheses. Significance levels: *** 1%, ** 5%, * 10%. Instrument: kindergarten teacher over kindergarten-aged (3-5 aged) population ratio is in the lowest sextile in the municipality the child lived at age 4. Missing variables imputed with the sample average and control variables are included of dummy-variables that denote missing imputed observations with 1.*

A.Table 14. OLS and IV estimates of the effect of an additional year spent in kindergarten on test score; various cutoffs for the instrument

Panel A	**mathematics**	**reading**	Panel B: 1st stage: **Years spent in kindergarten** **6th grade**
OLS	0.0613*** [0.004]	0.0594*** [0.004]	-
IV: teachers/children aged 3-5 < **33.** pctile	0.3312*** [0.050]	0.3937*** [0.050]	-0.1527*** [0.006]
IV: teachers/children aged 3-5 < **25**. pctile	0.2986*** [0.044]	0.3341*** [0.045]	-0.1836*** [0.007]
IV: teachers/children aged 3-5 < **20**. pctile	0.2052*** [0.044]	0.2467*** [0.045]	-0.1917*** [0.008]
IV: teachers/children aged 3-5 < **16.** pctile	0.1844*** [0.044]	0.2153*** [0.044]	-0.2084*** [0.008]
IV: teachers/children aged 3-5 < **14.** pctile	0.1563*** [0.042]	0.2012*** [0.042]	-0.2300*** [0.009]
IV: teachers/children aged 3-5 < **12.** pctile	0.1392*** [0.042]	0.1735*** [0.041]	-0.2422*** [0.009]
IV: teachers/children aged 3-5 < **10.** pctile	0.1640*** [0.044]	0.1795*** [0.044]	-0.2533*** [0.010]

*Notes: Own calculations. Data source: HNABC 2008, KIR-Stat 1996-2001, DEM 1996-2001. Source of the figures: Table15, Tabl16 and Table17 in the Appendix. Children with special needs excluded. All control variables listed in Table3 in the Appendix included. Scores are measured in standard deviations. Robust standard errors in parentheses. Significance levels: *** 1%, ** 5%, * 10%. Instrument: Scarcity of kindergarten teachers that equals one if kindergarten teachers over the 3-5 aged population ratio was lower than the appropriate cutoff-value in the municipality at the child's age of 4 in the HNABC data base.*

A.Table 15. IV estimates of the effect of time spent in kindergarten on test score; 6th grade, mathematics and reading, instrument with various cut-offs / I.

	33. percentile		25. percentile		20. percentile	
	mathematics	reading	mathematics	reading	mathematics	reading
Years spent in kindergarten	0.3312***	0.3937***	0.2986***	0.3341***	0.2052***	0.2467***
	[0.050]	[0.050]	[0.044]	[0.045]	[0.044]	[0.045]
Parental education 0-8th grade	-0.5619***	-0.5505***	-0.5707***	-0.5667***	-0.5963***	-0.5906***
	[0.019]	[0.020]	[0.018]	[0.018]	[0.018]	[0.018]
Parental education vocational	-0.3160***	-0.2894***	-0.3176***	-0.2925***	-0.3224***	-0.2969***
	[0.009]	[0.009]	[0.008]	[0.008]	[0.008]	[0.008]
Parental education tertiary	0.3539***	0.3304***	0.3538***	0.3301***	0.3533***	0.3297***
	[0.008]	[0.008]	[0.008]	[0.008]	[0.008]	[0.008]
Settlement type: County center	0.0399***	0.0211	0.0404***	0.0220	0.0419***	0.0234
	[0.015]	[0.015]	[0.015]	[0.015]	[0.015]	[0.015]
Settlement type: Town	-0.0801***	-0.0751***	-0.0801***	-0.0751***	-0.0800***	-0.0751***
	[0.013]	[0.013]	[0.013]	[0.013]	[0.013]	[0.012]
Settlement type: Village	-0.0877***	-0.1111***	-0.0883***	-0.1122***	-0.0901***	-0.1139***
	[0.014]	[0.014]	[0.014]	[0.013]	[0.014]	[0.013]
Region: Central Transdanubia	-0.0670***	-0.0452***	-0.0640***	-0.0398***	-0.0554***	-0.0318**
	[0.014]	[0.014]	[0.013]	[0.013]	[0.013]	[0.013]
Region: Western Transdanubia	-0.0072	0.0071	-0.0050	0.0111	0.0014	0.0171
	[0.014]	[0.013]	[0.013]	[0.013]	[0.013]	[0.013]
Region: Southern Transdanubia	-0.0527***	-0.0430***	-0.0482***	-0.0347**	-0.0353**	-0.0226
	[0.015]	[0.015]	[0.015]	[0.015]	[0.015]	[0.015]
Region: Northern Hungary	-0.0815***	-0.1044***	-0.0822***	-0.1058***	-0.0843***	-0.1077***
	[0.013]	[0.013]	[0.013]	[0.013]	[0.012]	[0.012]
Region: Northern Great Plain	-0.0334***	-0.0561***	-0.0324***	-0.0543***	-0.0296**	-0.0517***
	[0.012]	[0.012]	[0.012]	[0.012]	[0.012]	[0.012]
Region: Southern Great Plain	-0.0624***	-0.0521***	-0.0587***	-0.0453***	-0.0480***	-0.0353***
	[0.014]	[0.014]	[0.013]	[0.013]	[0.013]	[0.013]
Father employed permanently	0.0767***	0.0724***	0.0792***	0.0770***	0.0864***	0.0838***
	[0.011]	[0.011]	[0.010]	[0.010]	[0.010]	[0.010]
Primary school starting age	-0.2319***	-0.2426***	-0.2193***	-0.2194***	-0.1829***	-0.1855***
	[0.020]	[0.020]	[0.018]	[0.018]	[0.018]	[0.018]
Disadvantaged child	-0.0857***	-0.0789***	-0.0883***	-0.0837***	-0.0958***	-0.0906***
	[0.010]	[0.010]	[0.009]	[0.009]	[0.009]	[0.009]
Household size	-0.0051	-0.0207***	-0.0067**	-0.0236***	-0.0113***	-0.0280***
	[0.004]	[0.004]	[0.003]	[0.003]	[0.003]	[0.003]
Lives with biological mother	0.0383*	0.0511**	0.0406*	0.0554**	0.0473**	0.0618***
	[0.022]	[0.023]	[0.021]	[0.023]	[0.021]	[0.022]
Lives with biological father	-0.0104	-0.0142	-0.0108	-0.0150	-0.0120	-0.0161*
	[0.010]	[0.010]	[0.009]	[0.009]	[0.009]	[0.009]
Lives with stepmother	-0.1643***	-0.1105***	-0.1654***	-0.1124***	-0.1683***	-0.1152***
	[0.030]	[0.032]	[0.030]	[0.032]	[0.030]	[0.031]
Lives with stepfather	-0.1093***	-0.0754***	-0.1074***	-0.0718***	-0.1018***	-0.0665***
	[0.013]	[0.013]	[0.013]	[0.013]	[0.013]	[0.013]
Books at home: less than 50	-0.2049***	-0.2537***	-0.2088***	-0.2610***	-0.2202***	-0.2716***
	[0.012]	[0.012]	[0.012]	[0.012]	[0.012]	[0.012]
Internet at home	0.1240***	0.1169***	0.1267***	0.1218***	0.1343***	0.1288***
	[0.009]	[0.009]	[0.009]	[0.009]	[0.009]	[0.009]
Child works at home permanently	-0.1332***	-0.1615***	-0.1347***	-0.1642***	-0.1391***	-0.1683***
	[0.010]	[0.010]	[0.010]	[0.010]	[0.010]	[0.010]
Constant	0.6063***	0.5578***	0.6298***	0.6008***	0.6971***	0.6639***
	[0.062]	[0.062]	[0.059]	[0.060]	[0.059]	[0.059]
dummies for imputed missings	yes	yes	yes	yes	yes	yes
Observations	79,507	79,506	79,507	79,506	79,507	79,506
R-squared	0.212	0.209	0.221	0.228	0.239	0.249

*Notes: Own calculations. Data source: HNABC 2008, KIR-Stat 1996-2001, DEM 1996-2001. Children with special needs excluded. Reference categories: Parental education: secondary, Settlement type: Budapest, Region: Central-Hungary. Scores are measured in standard deviations. Robust standard errors in parentheses. Significance levels: *** 1%, ** 5%, * 10%. Instrument: kindergarten teacher over kindergarten-aged (3-5 aged) population ratio is lower than the appropriate cutoff-value in the municipality the child lived at age 4. Missing variables imputed with the sample average and control variables are included of dummy-variables that denote missing imputed observations with 1.*

A.Table 16. IV estimates of the effect of time spent in kindergarten on test score; 6th grade, mathematics and reading, instrument with various cut-offs / II.

	14. percentile		12. percentile		10. percentile	
	mathematics	reading	mathematics	reading	mathematics	reading
Years spent in kindergarten	0.1563*** [0.042]	0.2012*** [0.042]	0.1392*** [0.042]	0.1735*** [0.041]	0.1640*** [0.044]	0.1795*** [0.044]
Parental education 0-8th grade	-0.6096*** [0.017]	-0.6030*** [0.018]	-0.6143*** [0.017]	-0.6106*** [0.018]	-0.6075*** [0.018]	-0.6089*** [0.018]
Parental education vocational	-0.3249*** [0.008]	-0.2993*** [0.008]	-0.3258*** [0.008]	-0.3007*** [0.008]	-0.3245*** [0.008]	-0.3004*** [0.008]
Parental education tertiary	0.3531*** [0.008]	0.3295*** [0.008]	0.3530*** [0.008]	0.3294*** [0.008]	0.3531*** [0.008]	0.3294*** [0.008]
Settlement type: County center	0.0427*** [0.015]	0.0242* [0.015]	0.0430*** [0.015]	0.0246* [0.015]	0.0426*** [0.015]	0.0245* [0.015]
Settlement type: Town	-0.0800*** [0.013]	-0.0751*** [0.012]	-0.0800*** [0.013]	-0.0751*** [0.012]	-0.0800*** [0.013]	-0.0751*** [0.012]
Settlement type: Village	-0.0910*** [0.014]	-0.1148*** [0.013]	-0.0913*** [0.014]	-0.1153*** [0.013]	-0.0908*** [0.014]	-0.1152*** [0.013]
Region: Central Transdanubia	-0.0510*** [0.013]	-0.0276** [0.013]	-0.0494*** [0.013]	-0.0250* [0.013]	-0.0517*** [0.013]	-0.0256** [0.013]
Region: Western Transdanubia	0.0047 [0.013]	0.0202 [0.013]	0.0059 [0.013]	0.0221* [0.013]	0.0042 [0.013]	0.0217* [0.013]
Region: Southern Transdanubia	-0.0285** [0.014]	-0.0163 [0.014]	-0.0262* [0.014]	-0.0125 [0.014]	-0.0296** [0.015]	-0.0133 [0.014]
Region: Northern Hungary	-0.0854*** [0.012]	-0.1087*** [0.012]	-0.0858*** [0.012]	-0.1093*** [0.012]	-0.0853*** [0.012]	-0.1092*** [0.012]
Region: Northern Great Plain	-0.0282** [0.012]	-0.0504*** [0.012]	-0.0277** [0.012]	-0.0495*** [0.012]	-0.0284** [0.012]	-0.0497*** [0.012]
Region: Southern Great Plain	-0.0424*** [0.013]	-0.0301** [0.013]	-0.0405*** [0.013]	-0.0269** [0.013]	-0.0433*** [0.013]	-0.0276** [0.013]
Father employed permanently	0.0902*** [0.010]	0.0873*** [0.010]	0.0915*** [0.010]	0.0894*** [0.010]	0.0896*** [0.010]	0.0890*** [0.010]
Primary school starting age	-0.1639*** [0.017]	-0.1677*** [0.017]	-0.1573*** [0.017]	-0.1570*** [0.017]	-0.1669*** [0.018]	-0.1593*** [0.018]
Disadvantaged child	-0.0997*** [0.009]	-0.0942*** [0.009]	-0.1010*** [0.009]	-0.0964*** [0.009]	-0.0990*** [0.009]	-0.0960*** [0.009]
Household size	-0.0138*** [0.003]	-0.0302*** [0.003]	-0.0146*** [0.003]	-0.0316*** [0.003]	-0.0134*** [0.003]	-0.0313*** [0.003]
Lives with biological mother	0.0509** [0.021]	0.0651*** [0.022]	0.0521** [0.021]	0.0671*** [0.022]	0.0503** [0.021]	0.0667*** [0.022]
Lives with biological father	-0.0127 [0.009]	-0.0167* [0.009]	-0.0129 [0.009]	-0.0170* [0.009]	-0.0126 [0.009]	-0.0169* [0.009]
Lives with stepmother	-0.1699*** [0.030]	-0.1166*** [0.031]	-0.1705*** [0.029]	-0.1175*** [0.031]	-0.1697*** [0.030]	-0.1173*** [0.031]
Lives with stepfather	-0.0988*** [0.013]	-0.0637*** [0.013]	-0.0978*** [0.013]	-0.0621*** [0.013]	-0.0993*** [0.013]	-0.0624*** [0.013]
Books at home: less than 50	-0.2262*** [0.011]	-0.2772*** [0.012]	-0.2283*** [0.011]	-0.2806*** [0.011]	-0.2252*** [0.012]	-0.2798*** [0.012]
Internet at home	0.1382*** [0.008]	0.1325*** [0.009]	0.1396*** [0.008]	0.1348*** [0.008]	0.1376*** [0.009]	0.1343*** [0.009]
Child works at home permanently	-0.1414*** [0.010]	-0.1704*** [0.010]	-0.1422*** [0.010]	-0.1717*** [0.010]	-0.1410*** [0.010]	-0.1714*** [0.010]
Constant	0.7324*** [0.058]	0.6967*** [0.058]	0.7447*** [0.057]	0.7167*** [0.058]	0.7268*** [0.058]	0.7124*** [0.058]
dummies for imputed missings	yes	yes	yes	yes	yes	yes
Observations	79,507	79,506	79,507	79,506	79,507	79,506
R-squared	0.245	0.256	0.246	0.260	0.244	0.259

*Notes: Own calculations. Data source: HNABC 2008, KIR-Stat 1996-2001, DEM 1996-2001. Children with special needs excluded. Reference categories: Parental education: secondary, Settlement type: Budapest, Region: Central-Hungary. Scores are measured in standard deviations. Robust standard errors in parentheses. Significance levels: *** 1%, ** 5%, * 10%. Instrument: kindergarten teacher over kindergarten-aged (3-5 aged) population ratio is lower than the appropriate cutoff-value in the municipality the child lived at age 4. Missing variables imputed with the sample average and control variables are included of dummy-variables that denote missing imputed observations with 1.*

A.Table 17. IV 1st stage estimates on years spent in kindergarten, 6th grade, instrument with different cutoff-values

	Years spent in kindergarten					
	33.pct	25. pct.	20. pct	14. pct	12. pct	10. pct
instrument with appropriate cutoff	-0.1527***	-0.1836***	-0.1917***	-0.2300***	-0.2422***	-0.2533***
	[0.006]	[0.007]	[0.008]	[0.009]	[0.009]	[0.010]
Parental education 0-8th grade	-0.2629***	-0.2612***	-0.2614***	-0.2622***	-0.2614***	-0.2639***
	[0.013]	[0.013]	[0.013]	[0.013]	[0.013]	[0.013]
Parental education vocational	-0.0483***	-0.0483***	-0.0483***	-0.0487***	-0.0488***	-0.0490***
	[0.006]	[0.006]	[0.006]	[0.006]	[0.006]	[0.006]
Parental education tertiary	-0.0073	-0.0065	-0.0064	-0.0055	-0.0054	-0.0045
	[0.006]	[0.006]	[0.006]	[0.006]	[0.006]	[0.006]
Settlement type: County center	0.0524***	0.0315**	0.0286**	0.0135	0.0141	0.0103
	[0.012]	[0.012]	[0.012]	[0.012]	[0.012]	[0.012]
Settlement type: Town	0.0631***	0.0508***	0.0387***	0.0157	0.0118	0.0037
	[0.011]	[0.011]	[0.011]	[0.011]	[0.011]	[0.011]
Settlement type: Village	0.0799***	0.0707***	0.0607***	0.0405***	0.0406***	0.0293***
	[0.012]	[0.012]	[0.012]	[0.011]	[0.011]	[0.011]
Region: Central Transdanubia	0.0541***	0.0652***	0.0682***	0.0821***	0.0833***	0.0859***
	[0.011]	[0.010]	[0.010]	[0.010]	[0.010]	[0.010]
Region: Western Transdanubia	0.0249**	0.0385***	0.0411***	0.0577***	0.0588***	0.0658***
	[0.011]	[0.011]	[0.011]	[0.010]	[0.010]	[0.010]
Region: Southern Transdanubia	0.0939***	0.1050***	0.1115***	0.1323***	0.1347***	0.1422***
	[0.011]	[0.011]	[0.011]	[0.011]	[0.011]	[0.011]
Region: Northern Hungary	-0.0427***	-0.0309***	-0.0262**	-0.0088	-0.0153	-0.0094
	[0.011]	[0.011]	[0.011]	[0.011]	[0.011]	[0.011]
Region: Northern Great Plain	0.0295***	0.0338***	0.0344***	0.0450***	0.0449***	0.0461***
	[0.010]	[0.010]	[0.010]	[0.010]	[0.010]	[0.010]
Region: Southern Great Plain	0.0944***	0.0911***	0.0921***	0.1069***	0.1060***	0.1078***
	[0.010]	[0.010]	[0.010]	[0.010]	[0.010]	[0.010]
Father employed permanently	0.0747***	0.0742***	0.0742***	0.0732***	0.0730***	0.0737***
	[0.009]	[0.009]	[0.009]	[0.009]	[0.009]	[0.009]
Primary school starting age	0.3951***	0.3918***	0.3920***	0.3911***	0.3899***	0.3894***
	[0.005]	[0.005]	[0.005]	[0.005]	[0.005]	[0.005]
Disadvantaged child	-0.0771***	-0.0756***	-0.0751***	-0.0739***	-0.0745***	-0.0738***
	[0.007]	[0.007]	[0.007]	[0.007]	[0.007]	[0.007]
Household size	-0.0482***	-0.0480***	-0.0479***	-0.0481***	-0.0481***	-0.0483***
	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]
Lives with biological mother	0.0720***	0.0736***	0.0749***	0.0743***	0.0731***	0.0733***
	[0.020]	[0.019]	[0.019]	[0.019]	[0.019]	[0.019]
Lives with biological father	-0.0101	-0.0103	-0.0109	-0.0104	-0.0103	-0.0106
	[0.008]	[0.008]	[0.008]	[0.008]	[0.008]	[0.008]
Lives with stepmother	-0.0298	-0.0291	-0.0289	-0.0315	-0.0340	-0.0315
	[0.028]	[0.028]	[0.028]	[0.028]	[0.028]	[0.028]
Lives with stepfather	0.0600***	0.0596***	0.0594***	0.0591***	0.0593***	0.0596***
	[0.010]	[0.010]	[0.010]	[0.010]	[0.010]	[0.010]
Books at home: less than 50	-0.1174***	-0.1163***	-0.1173***	-0.1174***	-0.1172***	-0.1177***
	[0.010]	[0.010]	[0.010]	[0.010]	[0.010]	[0.010]
Internet at home	0.0755***	0.0747***	0.0744***	0.0744***	0.0741***	0.0746***
	[0.006]	[0.006]	[0.006]	[0.006]	[0.006]	[0.006]
Child works at home permanently	-0.0423***	-0.0417***	-0.0412***	-0.0410***	-0.0414***	-0.0424***
	[0.008]	[0.008]	[0.008]	[0.008]	[0.008]	[0.008]
Constant	0.6783***	0.6992***	0.6969***	0.7044***	0.7137***	0.7168***
	[0.041]	[0.041]	[0.041]	[0.041]	[0.041]	[0.041]
dummies for imputed missings	yes	yes	yes	yes	yes	yes
Observations	79,531	79,531	79,531	79,531	79,531	79,531
R-squared	0.138	0.140	0.139	0.140	0.141	0.140

*Notes: Own calculations. Data source: HNABC 2008, KIR-Stat 1996-2001, DEM 1996-2001. Children with special needs excluded. Reference categories: Parental education: secondary, Settlement type: Budapest, Region: Central-Hungary. Robust standard errors in parentheses. Significance levels: *** 1%, ** 5%, * 10%. Instrument: kindergarten teacher over kindergarten-aged (3-5 aged) population ratio is lower than the appropriate cutoff-value in the municipality the child lived at age 4. Missing variables imputed with the sample average and control variables are included of dummy-variables that denote missing imputed observations with 1.*

Reference List

Angrist, J.D., Imbens, G.W. 1995. "Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity". *Journal of the American Statistical Association,* 90(430), pp. 431–442.

Angrist, J.D., Imbens, G.W., Rubin, D.B. 1996. "Identification of Causal Effects Using Instrumental Variables". *Journal of the American Statistical Association*, 91(434), pp. 444-455.

Angrist, J.D., Krueger, A.B. 1999. "Empirical Strategies in Labor Economics". In: Ashenfelter, O. & Card, D. (ed.), 1999. "*The Handbook of Labor Economics*", Elsevier, Vol.3, pp. 1277-1366

Angrist, J.D., Krueger, A.B. 2001. „Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments*". Journal of Economic Perspectives*, 15(4), pp. 69-85.

Bayer, P. 2000. "Tiebout Sorting and Discrete Choices: A New Explanation for Socioeconomic Differences in the Consumption of School Quality". *Unpublished Manuscript, Stanford University,* available at: http://ase.tufts.edu/econ/papers/patrick_bayer.pdf, downloaded: 9th May, 2011

Berlinski, S., Galiani, S., Gertler, P. 2009. "The effect of pre-primary education on primary school performance". *Journal of Public Economics*, 93 (1) pp. 219–234.

Björklund, A., Moffitt, R. 1987. "The Estimation of Wage Gains and Welfare Gains in Self-Selection". *The Review of Economics and Statistics,* 69(1), pp. 42-49.

Blau, D., Currie, J., 2006. "Pre-school, day care, and after school care: who's minding the kids?". In: Welch, F. and Hanushek, E. (ed.), 2006. "*The Handbook of the Economics of Education"*, Elsevier, Vol. 1, pp. 1163–1278.

Bound, J., Jaeger, D.A., Baker, R.M. 1995. "Problems with Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogeneous Explanatory Variable is Weak". J*ournal of the American Statistical Association,* 90(430), pp. 443-450.

Brooks-Gunn, J. 2003. "Do You Believe In Magic? What We Can Expect From Early Childhood Intervention Programs". *Social Policy Report*, 18(1)

Card, D. 1995. "Earnings, Schooling and Ability Revisited". *NBER Working Paper No. 4832*

Card, D. 1999. "The Causal Effect of Education on Earnings". In: Ashenfelter, O. & Card, D. (ed.), 1999. "The *Handbook of Labor Economics*", Elsevier, Vol.3, pp. 1801-1863.

Card, D. 2001. "Estimating the Returns to Schooling: Progress on Some Persistent Econometric Problems". *Econometrica,* 69 (5), pp. 1127-1160.

Chetty, R., Friedman, J.N., Hilger, N., Saez, E., Schanzenbach, D.W., Yagan, D. 2011. "How does your Kindergarten Classroom affect your Earnings? Evidence from Project Star". Mimeo, available: http://obs.rc.fas.harvard.edu/chetty/STAR.pdf, downloaded: 22th May, 2011

Cunha, F., Heckman, J. J., Lochner, L. and Masterov, D. V. 2006. "Interpreting the Evidence on Life Cycle Skill Formation". In: Welch, F. and Hanushek, E. (ed.), 2006. "*The Handbook of the Economics of Education"*, Elsevier, Vol. 1, pp. 697-812.

Cunha, F., Heckman, J. J. 2007. „The Technology of Skill Formation". *NBER Working Paper No.12840*

Currie, J. 2001. "Early Childhood Education Programs". *Journal of Economic Perspectives,* 15(1), pp. 213-238.

Hanushek, E. A., Kain, J. F., Rivkin, S. G. 1998. "Teachers, Schools, and Academic Achievement". *NBER Working Paper, No. 6691.*

Hanushek, E. A. 2002. "Publicly provided education". *NBER Working Paper No. 8799*

Havas, G. 2008. "Equality of opportunity, desegregation". In Fazekas, K., Köllő, J. & Varga, J. (ed.). 2008. "*Green Book for the Renewal of Public Education in Hungary*", Ecostat, Budapest, pp. 131-149.

Heckman, J. J., Carneiro, P. 2003. "Human Capital Policy". *NBER Working Paper No. 9495*

Heckman, J.J., Lochner, L.J., Todd, P.E. 2006a. "Earning Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond". In: Hanushek, E. & Welch, F. (ed.), 2006. *"The Handbook of the Economics of Education*", Elsevier, Vol.1, pp. 307-458.

Heckman, J.J., Urzua, S., Vytlacil, E. 2006b. "Understanding Instrumental Variable in Models with Essential Heterogeneity". *The Review of Economics and Statistics*, 88(3), pp. 389-432.

Heckman, J.J. 2007. "The Economics, Technology, and Neuroscience of Human Capability Formation". *Proceedings of the National Academy of Sciences of the United States of America,* 104 (33), pp. 13250-13255.

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

Imbens, G.W., Angrist, J.D. 1994. "Identification and Estimation of Local Average Treatment Effects". *Econometrica,* 62(2), pp. 467-475.

Katz, L.F., Autor, D.H. 1999. "Changes in the wage structure and earnings inequality". In: Ashenfelter, O. & Card, D. (ed.), 1999. "*Handbook of Labor Economics*", Volume 3, Elsevier, pp. 1463-1555.

Ludwig, J., Miller, D.L., 2007. "Does Head Start improve children's life chances? Evidence from a regression discontinuity design". *Quarterly Journal of Economics,* 122(1), pp. 159–208.

Magnuson, K.A., Ruhm, C., Waldfogel, J. 2007. "Does prekindergarten improve school preparation and performance?" *Economics of Education Review*, 26(1), pp. 33-51.

Rivers, J. C., Sanders, W. L. 2002. "Teacher Quality and Equity in Educational Opportunity: Findings and Policy Implications". In: Izumi, L .T. & Eders, W. M. (ed.), 2002. "*Teacher Quality*". Hoover Institution Press, Stanford, CA. pp. 13-24.

Shonkoff, J. 2003. "From Neurons to Neighborhoods: Old and New Challenges for Developmental and Behavioral Pediatrics". *Journal of Developmental & Behavioral Pediatrics*, 24(1), pp. 70-76.

Todd, P.E., Wolpin, K.I. 2007. "The Production of Cognitive Achievement in Children: Home, School, and Racial Test Score Gaps". *Journal of Human Capital,* 1(1), pp. 91-136.

Willis, R. J. 1987. "Wage determinants: A survey and reinterpretation of human capital earnings functions". In: Ashenfelter, O. & Layard, R. (ed.), 1987. "*The Handbook of Labor Economics*". Elsevier, Vol.1, pp. 525-602.

Wooldridge, J.M. 1999. "*Introductory Econometrics. A Modern Approach".* 2nd edition. Cambridge, MIT Press.