Eötvös Loránd University

Faculty of Social Sciences

BACHELOR OF ART PROGRAM

The Effects of Parental Separation on Children’s Academic Performance: *Empirical Evidence from Hungary*

Supervisor: Author:

Horn Dániel Nguyen Nam Son

GES0XD

Applied Economics

April 2021

Table of Contents

[Chapter 1. - Introduction 4](#_Toc64591242)

[1.1. Purpose and Scope 4](#_Toc64591243)

[1.2. Demographic Overview of Hungary 6](#_Toc64591244)

[1.2.1. Divorce Rate 6](#_Toc64591245)

[1.2.2. Children at the Separation 6](#_Toc64591246)

[1.2.3. Single-Parent Families 7](#_Toc64591247)

[1.2.4. Remarriage 8](#_Toc64591248)

[Chapter 2. - Literature Review 8](#_Toc64591249)

[2.1. Independent Variables 8](#_Toc64591250)

[2.1.1. Single-Parent Families 8](#_Toc64591251)

[2.1.2. Stepfamilies 9](#_Toc64591252)

[2.2. Intermediary Variables 10](#_Toc64591253)

[2.2.1. Economic Resources 10](#_Toc64591254)

[2.2.2. Parental Qualifications 11](#_Toc64591255)

[2.2.3. Parental School Involvement (PSI) and Home Environment 12](#_Toc64591256)

[2.2.4. Residential Mobility 13](#_Toc64591257)

[2.3. Dependent variables 15](#_Toc64591258)

[2.3.1. Grades and Test Scores 15](#_Toc64591259)

[2.3.2. Classroom Behavior 15](#_Toc64591260)

[2.3.3. Expel and Suspension 16](#_Toc64591261)

[2.3.4. Grade Repeat and Dropping Out 16](#_Toc64591262)

[2.3.5. Educational Attainment 18](#_Toc64591263)

[2.4. Control Variables 18](#_Toc64591264)

[2.4.1. Age Differences 18](#_Toc64591265)

[2.4.2. Gender Differences 19](#_Toc64591266)

[2.4.3. Ethnic and Racial Differences 20](#_Toc64591267)

[2.4.4. Birthweight and Health Status 21](#_Toc64591268)

[2.4.5. School Quality 21](#_Toc64591269)

[Chapter 3. - Data Source and Methods 22](#_Toc64591270)

[3.1. Frequencies and Cross Tabulations 22](#_Toc64591271)

[3.2. Central Tendencies 23](#_Toc64591272)

[3.3. Feature selection 25](#_Toc64591273)

[3.4. Methods 26](#_Toc64591274)

[3.4.1. Value-Added Model (VAM) 26](#_Toc64591275)

[3.4.2. Staggered Difference in Differences 29](#_Toc64591276)

[Chapter 4. Results 37](#_Toc64591277)

[4.1. Value-Added Model (VAM) Estimation 37](#_Toc64591278)

[4.2. Staggered DID Estimations 39](#_Toc64591279)

[4.2.1. Upper-bound Estimates 39](#_Toc64591280)

[4.2.2. Lower-bound Estimates 52](#_Toc64591281)

[4.2.3. Additional Results 54](#_Toc64591282)

[Chapter 5. Limitations 56](#_Toc64591283)

[Chapter 6. Conclusion 58](#_Toc64591284)

[References 59](#_Toc64591285)

[Appendix 65](#_Toc64591286)

# Chapter 1. - Introduction

## 1.1. Purpose and Scope

Since the mid-1960s scholars have been studying the risks associated with parental separation on children’s academic grades. Having been exposed to the in-household conflict leading to the absence of a parent, children tend to perform less well at school. The purpose of this thesis is to empirically support the distinguished results of existing literature and to provide a bird’s-eye view of the adverse effects of parental separation on educational outcomes in Hungary. The Difference-in-differences (DID) is the most common approach used to evaluate intervention results, but instead of the canonical two-by-two format (two-way fixed effect), we will first estimate a static value-added model (VAM) before turning to a more complex version of DID with multiple groups and time periods (staggered adoption). The former approximation will regress the 2009 grade point averages (GPA on a Hungarian scale of 1 to 5) on that of 2006, intermediary, and sets of control variables which may contribute to the effect of disruption. The latter will require the comparison of average robust differences between groups and non-intact and intact families after the separation (difference in differences) accounting for intermediary channels that potentially explain the gap in scores. For the empirical purpose, I will invoke repeated cross-sectional data (approx. 10 to 8 thousand respondents in each year from 2006 to 2009) extracted from the Hungarian Life Course Survey (HLCS) of TÁRKI which followed a cohort of adolescents between 2006 and 2012. Such a four-year time period should be able to cover the academic performances ranging from 8th and 12th grade of secondary school.

There are several arguments, why my essay should contribute to the existing pool of knowledge in such a matter. The dataset at-hand offers the luxury to browse many more characteristics of the respondents compared to the so-called “competence assessment” in Hungary, therefore the issue of heterogeneity is treated in the majority of the cases. Also, we can explain more variances in the measures of performance and control for tons of differences to arrive at the pure impact. Besides, the competence assessment is only conducted every two years from 6th grade to 10th grade, whereas the Hungarian Life Course Survey (HLCS) dataset has a wider range of observations and a year-to-year assessment covering the whole span of secondary education. Additionally, the major share of the papers related to the issue does not account for all the intermediary variables between separation and performance, while my paper will touch upon most of the key determinants collected from a diverse pool of studies. Furthermore, while the VAM will provide a birds’ eye view of the matter, the implemented estimation method of staggered DID offers much more flexibility and causal inferential power than the mere observation of DID coefficient in OLS regressions. The approach implemented in this essay accounts for all the known caveats of the DID, therefore my results are more representative than that of previous literature. Finally, the empirical discussion of such phenomena in Hungary is very limited, due to the lack of literature, thus, I expect to shed light on the issue specifically and encourage parents, teachers, and policymakers to consider my findings when designing policies to target such issue.

This paper can be divided into six main sections. In the first chapter, I will demonstrate a brief demographic overview of the children of parental separation in Hungary. In Chapter 2., I will summarize a vast body of relevant literature that discusses the underlying phenomena between family structure and academic performance, while identifying the key contributors to our model. Chapter 3. will turn to the empirical approach of the VAM and Staggered DID which will estimate the effect of family disruption on GPA, appended with an assessment on the math grades based on the forward selection methodology and the literature review. In Chapter 4, I will discuss the results of both types of estimation, and choose the best models for estimation based on applicable criteria. Lastly, in the remaining parts, Chapter 5. and 6., I will introduce the caveats and limitations of this study and reprise the key takeaways of my findings.

Since the formal and data constraints do not allow me to empirically discuss the issue from every aspect, hence, I will put more emphasis on initial parental separations rather than on remarriage, and other blended and extreme family types. Stated differently, only the differences between intact and non-intact families will appear at the literature and empirical level, however, the reader should be provided with adequate literature to explore the topic in detail. Also, I will distinguish parental divorce (legally) from separation, therefore, the determinant of treatment will be based on the question of whether the child stays in the same household with both of their parents.

## 1.2. Demographic Overview of Hungary

### 1.2.1. Divorce Rate

In the mid-1970s, the crude marriage rate (CMR) (per 1000 people) in Hungary was four times higher than the crude divorce rate (CDR). Since then, the number of marriages had been in decline and the divorce rate remained constant around 20-25 thousand per year, so that between 2001 and 2011 the gap between the two rates significantly decreased, as the marriage rate is only 1.7 times higher than the divorce rate. Thanks to the recovery of the marriage rate due to the better conditions, the ratio between the two rates was 2.6 in 2016 (Központi Statisztikai Hivatal, 2018). The most common length of marriage duration in Hungary and the OECD countries is 20 or more years followed by the group of divorcees after 5 to 9 years (OECD, 2017).

### 1.2.2. Children at the Separation

Between 1990 and 2017, the ratio of divorces without underage children increased by nearly 10% (from 36.5% to 46.2%) and decreased among those with one or more underage children. Such an increase in the prior group can be explained by the general decline in fertility and the increase in the proportion of childless families, although it is also conceivable for couples to wait until their children reach adulthood and only then vouch for divorce. It is also supported by the fact that between 1990 and 2017, the average number of divorces and the number of divorces after 20 years of marriage increased (Központi Statisztikai Hivatal, 2019).

According to the latest micro census in 2016, 18% of children aged 18 or below and 26% of 17-18-year-olds were raised in a single-parent family at that time. The older the age group, the higher the value of the indicator. However, those who have already experienced the breakdown of their parents' relationship, or the loss of one parent, but now may live in blended families corresponds to a higher share. Such characteristic is backed by the fact that 31% of 0-18-year-olds do not live in an intact family, i. e. where at least one of the children living there does not grow up with both blood parents (Központi Statisztikai Hivatal, 2018; Monostori, 2019).

In terms of divorces associated with the number of “common” children, the most significant change was in the case of divorces with one underage child, as their proportion decreased from 36 to 30%. The decrease in the number of divorces with two underage children was less expressive (from 22% to 19%). Howbeit, the divorce rate for three or more underage children increased until the second half of the 2000s, and in the last ten years, it decreased significantly, by almost 30%, until it accounted for 4.9% of all divorces in 2017. Such patterns in the number of involved children are also present in the OECD averages (Központi Statisztikai Hivatal, 2019; OECD, 2017). To conclude, divorces in 2017 imposed negative effects on approximately 15.5 thousand children, compared to about 24.5 thousand ten years earlier (Központi Statisztikai Hivatal, 2019).

### 1.2.3. Single-Parent Families

In Hungary, most of the single-parent families stem from either the death of a parent (widowhood) or divorce. Within the groups of families, the ratio of such types of families was 10% until 1980. The proportion increased from 16% in 1990 to 20% two decades later, then it corrected to a level where nearly every fifth, as 18.3% of the families were single-parent families in 2016 (Központi Statisztikai Hivatal, 2018). In comparison, the share of single-parent households in the OECD countries ranges between 5-10% (OECD, 2016).

According to the population census in 2011, 1.315 million parents and children were living in single-parent families. Among those families, the children were often brought up and nurtured by the mother as the share of single-mother children's families was close to 86%, and single-mother families with children aged under 15 constituted 91% (Központi Statisztikai Hivatal, 2018; Monostori & Murinkó, 2015).

In 2016, the number of families with two married parents has been 1.757 million which accounts for 64% of the whole family population, while the number of single-parent families has stopped increasing in 2011 and then increased by 34 thousand throughout the five years until 503 thousand families. Within the groups of families with children, 56% of the families were based on marriage, and 30% of them were single mother or father families. Such an increase was due to the high number of divorces, the instability of cohabiting unions, and the postponement of adolescents moving out of the parental home. Only in every seventh single-parent family, the father was living with their children. In 2016, 35% of the families were consisting of married couples with children. Such a ratio has been decreasing since 2001, as it was 45% and 38% in 2011. (Judit & Lívia, 2015; Központi Statisztikai Hivatal, 2018, 2019).

As in the prior surveys, the territorial characteristics of the family structure remained intact throughout the years as the ratio of single-parent families was the highest in Budapest, accounting for 20% of the families. If we observe the ratio for each settlement status, we can notice that the weaker the territorial status, the smaller the share of single-parent families. (Központi Statisztikai Hivatal, 2018).

### 1.2.4. Remarriage

Between 1990 and 2017, in the vast majority of divorces (84%), the parties divorced their first marriage, i.e., they were single or unmarried before they got married. However, the proportion of unmarried men is lower each year than that of unmarried women. This is explained by the share of previously divorced marital status as they remarried and divorced again. Figuratively speaking, 14 to 17% of the women and 15 to 18% of men were divorced previously in the years studied. Such ratio has decreased compared to 1990, but since the turn of the millennium, the share of divorces where one of the parties has already divorced once has increased in 2017. Among men, these rates were higher than for women in each of the years studied, so it can be concluded that men remarry more often, but also divorce more times than women (Központi Statisztikai Hivatal, 2019).

# Chapter 2. - Literature Review

## 2.1. Independent Variables

### 2.1.1. Single-Parent Families

Numerous prior and ongoing studies have shown that there exists a recognizable relationship between parental separation and the children’s academic performance. Children of family disruption may face ongoing parental conflict beforehand, and socioeconomic disadvantage, as well as other stressors after the breakdown of the biological family (Shinoda, 2011). These impacts altogether have reportedly affected not only the children’s performance at school but also their classroom behavior and educational aspirations.

It has been not discussed in detail whether parental divorce or death has larger negative effects on children's academic achievement. Like divorce, parental death can also impose an adverse effect on children. What mainly distinguishes the parental death from divorce is that it is unprecedented therefore there is no conflict or prior tension as in the case of divorce. Those children who are in an intact family with high parental conflict tend to adjust relatively slower than those who are in a low-conflict divorced family (Amato & Anthony, 2014; Mechanic & Hansell, 1989). This suggests that parental conflict can be more harmful than the changing structure of the family. However, by being an observer of the continuous conflict and tension, children can prepare themselves even for the worst-case scenario and may experience milder effects in the short-run after the separation.

Comparing the deceased and divorced single-mother families, we have evidence that children in divorced family types tend to achieve lower levels of education. On the contrary, children from widowed single-mother households have a lower likelihood of completing secondary education. Although both types of stepfamilies emerging from either death or divorce achieved less in school on average, Biblarz (2016) argues that the disadvantages of widowhood in comparison to parental divorce can be offset by the remarriage, while a remarriage after a divorce is ambiguous. (Biblarz et al., 2016).

A study using growth-curve models asserted that children who have not yet passed the pubertal apex and are part of stable families[[1]](#footnote-1) have somewhat lower scores than that of the always-married families. In the meantime, children from unstable families[[2]](#footnote-2) performed worse than stable families compared to the biological family. It is needed to be mentioned that the differences between the groups were rather modest negative effects, but over time they have increased. Besides, the stabilized structure does not reduce the cumulative effect of the initial divorce because the slope of grade changes is minuscule, but rather prevents the children from performing at even lower levels. The findings suggest that the most affected family type was the unstable family and although the stability of the family does not mitigate the initial negative effect, the size of the damage is less than half of such unstable families, which could make children avoid the less positive trajectories (Yongmin Sun & Yuanzhang Li, 2009).

### 2.1.2. Stepfamilies

Although my empirics are not considering stepfamilies specifically, I find it important to discuss the relevant findings in short. Jeynes (2000) reported that when controlling for all the relevant variables (e.g., SES, race, gender, etc.) the largest negative effect on children’s academic achievement is associated with the widowed but remarried families, followed by the cohabiting households (when there is no legal bond between the parents). If the baseline model of no-SES variables is estimated, the absolute coefficients of family structure are greater which implies that the family structure variable is not adequate to estimate the nonbiased relationship between the academic performance and parental transition, and the family structure and SES variables are not correlated. As one would expect, the parental separation (either death or divorce) lowers the SES features of the family, while remarriage increases it due to the additional income input. However, the second transition may put extra downward pressure on the children. Also, children from the widowed remarried family are more likely to fall back and repeat a school grade than their counterparts who are in a widowed but single-parent family. Likewise, the likelihood of repeating a grade for children from divorced remarried families is larger than that of their peers in non-remarried divorced families. Thereby, single-parent and blended families are worse than two-parent biological families academic wise, while it is apparent that remarriage exerts additional adverse effects independently from the form of parental separation (Jeynes, 2000).

## 2.2. Intermediary Variables

### 2.2.1. Economic Resources

Generally, when one parent is no longer part of the household, their individual income is not (or partly) contributed to the children’s education and upbringing, therefore the economic resources of a single-parent family are much more constrained resulting in diminishing parental and financial investments. In other words, we are ought to consider income loss as one of the strongest intermediary variables between parental breakdown and the decline in academic performance.

Downey (1994) concluded that single fathers earn nearly twice the income of single mothers, which is consistent with the hypothesis and the previous literature concerning income. Hence, income seems to be a major key in the negative consequences associated with mother-only families, while such an effect is not identified in the remarried families. (Downey, 1994). The sensitivity of single-mothers can be explained by the ever-existing issue that women are socialized to have obligations to the family rather than being the “breadwinner” in the household. Such characteristics of society are even reflected in sexual discrimination in the workplace. The difference between the original and step or partner families is rather expressed in terms of the drop in quality of parental behavior. To summarize, single-mothers impose rather economic inequality on their children in the long-term, while mother-stepfamilies transmit disadvantages in parental support and emotional quality in adult life. (Thomson et al., 1994).

Teachman (1987) observed that not merely the family income but also the investment in children related to income have to be included as an economic parental resource. According to him, the presence of educationally related objects such as books, dictionaries, paintings, or a specific room purposed for studying has a direct and positive impact (independently from their ethnic background) on the child’s performance in terms of grades, the years of schooling the child will achieve, and his or her attendance rate in school as it “fosters academic skills, motivation, and orientation” (Teachman, 1987). On average, single fathers provide an averagely greater level of economic resources, for example, 26% of single-father families furnish computers compared to that of single-mother families which is 16%. Subsequently, 42% of father-only families saved up enough money for further studies of their children, while such ratio in the case of mother-only families was 31% (Downey, 1994). This implies that a considerably great share of single-mothers is unable to afford private lessons, educational equipment, books, electronics which facilitate the children’s academic achievements (Amato & Keith, 1991).

### 2.2.2. Parental Qualifications

A very recent study focused on heterogeneity in 2020 found that the negative relationship between parental divorce and GPA was stronger (0.24 points) among adolescents with more educated parents compared to that of the adolescents of less-educated parents. Adolescents in divorced families where the highest education level was either at a secondary school level or Bachelor’s level incline to perform worse by a moderate margin compared to disrupted families where the highest education was not higher than the basic level education (Nilsen et al., 2020).

It seems that the magnitude of the negative association between divorce and academic achievement was driven by maternal education level, as children with somewhat more educated mothers underperformed when controlled for income and paternal education levels. A statistically stronger reduction in GPA among the groups of divorced mothers can be observed where the highest education was either secondary, Bachelors, Masters or Ph.D. in comparison to their peers with basic-level education after adjusted for the socioeconomic measures. In other words, educated divorced mothers can’t transmit their knowledge and qualifications to their children, possibly due to the “double burden” of working in a high (demanding) position earned by their qualities and investing time in their children (Nilsen et al., 2020).

### 2.2.3. Parental School Involvement (PSI) and Home Environment

An American paper which observed a high-school sophomore cohort in 1986, found that the presence of both parents has positive effects on children’s school achievement, due to relatively more encouragement and help in school work compared to single-parent families (Astone & McLanahan, 1991). Reportedly, in single-parent and stepparent families, the educational aspiration and the financial commitment is lower, while there is less parental involvement in schoolwork. The custodial parent must bear the burden of work to compensate for income loss, thus, less time can be invested and devoted to their children and their school matters.

On one hand, due to the lack of time and the amount of stress, they tend to provide less monitoring, take disciplinary actions, and establish rules in the household (Amato, 1987a; Jeynes, 2005). On the other hand, some mothers implied that they feel their household more cohesive, and they can invest more time in their children since they don’t have to negotiate values and beliefs with the other parent (Amato, 1987a). Children in a one-parent family generally have more household chores to help out the single parent in need, therefore, it reduces their time spent on homework and preparing for classes (Amato, 1987b).

When it comes to disciplining the children, it appears that rather the absence of a father weakens parental authority over the children in mother-only households. However, some studies have contradicted such a hypothesis as they found that classroom behavior just as bad regardless of the sex of the custodial parent. Controlling for economic and interpersonal resources including the home equipment, it also seems that there is no difference in academic grades and most of the variance in adjustment can be explained by the socioeconomic factors (Downey, 1994).

It is also worth noting that generally, the father contributes less to childcare and parental relationships because as in the case of most countries, including Hungary, lion’s share of the children lives with their mother after the separation. Scholars understand that fathers have to adjust to the loss of shared residence which can lead to the withdrawal of parental involvement in the form of declines in visitations, less child support payment, and diminishing quality of co-parental involvement (Astone & McLanahan, 1991). It is also believed that the involvement of the father as a child-bearer is an important determiner in the adverse effect of separation, as those children who were able to keep their close relationship with both of their custodial and non-custodial parents were more likely to adjust better to the change of family structure (Madden-Derdich & Leonard, 2000).

As we could see in the previous two subsections, time and money are likely interchangeable variables since time must be devoted to making money which will convert to food, shelter, clothing, etc., and they spend money to buy time/investment such as childcare. Notwithstanding, the time for creating economic resources is shifted from the time one can invest in their children in terms of emotional support and parenting. While the single-mother struggles with financial distress, the fathers’ issue often comes down to the fact that they cannot offer adequate emotional support to their children in the post-separation period down (Thomson et al., 1994).

### 2.2.4. Residential Mobility

It has been a continuing dilemma of whether geographical mobility hurts academic achievements and adjustments through the change of school. A study by G. Ingersoll et al. (1989) observed the enrollment patterns in a public school found that the group of students who were less-mobile consistently performed better at school, while the negative impact decreases with an increase in the grade levels. (Ingersoll et al., 1989).

A vast amount of literature has shown that the downward change in family structure, can increase residential mobility, therefore it should be considered as a potential intermediate factor between family structure and academic performance. Rationally speaking, the reduction of family size infers switching to another, preferably smaller residents which suit better the financials or other reasons such as the proximity of the workplace. In other scenarios, the custodial parent needs to leave the current resident with their children which causes the switch. Likewise, the separation not only deprives the social capital of the family, but also that of the child which was based on the school acquaintances and the community in the prior neighborhood (Tucker & Long, 2018). As the single-parent remarries, children can be exposed again to moving but it does not have the same disruptive effect, because blended families are more likely to live in better neighborhoods conditions due to the income entrance of the step- or cohabiting parent (South et al., 1998).

Students who were forced to change schools are often missing educational materials and can experience a drop in grades. Furthermore, they suffer from information asymmetry as the school system, the quality of classes, and teachers are unknown to them in the new community. Thereby, they cannot gain full advantage of the resources and teachers tend to neglect those students which they do not know well, especially if there is a record for frequent changes between schools. Regarding the emotional resources, new students may feel socially isolated which also affects not only their classroom behavior but also their school achievements (Astone & McLanahan, 1994).

The results of Astone & McLanahan (1994) supported the hypothesis that mobile children naturally perform worse, and they stay longer in school due to family disruption. Other results have shown that residential mobility accounts for 18% of the difference in academic performance between single-parent and two-parent families, even though the reduction is not statistically significant. In the case of stepfamilies, the relationship is statistically significant and the residential mobility accounts for 29% of the difference (Astone & McLanahan, 1994). Tucker & Long (2018) reported a significant positive relationship between the initial movement, and the number of school problems, as those children who were not living with both of their original parents and moved more than once, had noticeably and statistically more problems compared to the never-moved peers (Tucker & Long, 2018). Per contra, there also exist indications that the family residential mobility does not have much effect on academic achievement, even though that the regression coefficients were consistent with the hypothesis set by Astone & McLanahan (1994) (Jeynes, 1999).

Despite everything in support of the hypothesis of residential mobility mentioned throughout this section, some argue that one may omit residential mobility as a variable from their model because it acts like a “catch up” variable that incorporates a lot of hidden (unobservable) factors like personal valuation, discrimination, diligence, etc. However, some results were statistically significant, therefore, the interpretation of residential mobility as a mild factor between family structure and education can be used as a tool of causal inference (Jeynes, 1999).

## 2.3. Dependent variables

### 2.3.1. Grades and Test Scores

To reiterate, adolescents who are part of a single-parent or stepparent household are prone to have lower GPAs, test scores, poorer attendance records, and more conflicts with school authorities (Astone & McLanahan, 1994). These effects are constituted in the enumerated channels of economic resources, rate of parental involvement, and residential mobility.

Even though the effects are modest, scholars reported that parental divorce was associated with lower scores in reading, math, interpersonal skills, self-control, while it increased the inter and externalizing problems. On the contrary, the loss of a parent also affects the children but lacked statistical significance due to the small sample size as it only explained the decline in math scores and the effect of internalization problems. Comparing the magnitude of significant coefficients, they argued that the size of the effect is roughly the same for parental loss and divorce, however, we should note that the study considered a kindergarten cohort which differs from our age group. (Amato & Anthony, 2014)

The magnitude of school problems across different types of disrupted families seems to be similar for each group. Grade-wise, remarriage does not worsen the grade difference between the non-intact families, as it is rather explained by the control variables. As regards the college expectations, the family structure differences among non-intact families were also not existential given the remaining covariates (Manning & Lamb, 2003).

### 2.3.2. Classroom Behavior

As far as delinquent acts within the school are concerned, remarried stepfather families have greater levels of misbehavior in comparison with single-mother families, but similar to that of the cohabiting stepfather families when we account for the intermediary factors, therefore the number of transitions determines the patterns of delinquency. (Manning & Lamb, 2003).

From the teachers’ perspective, they have also noticed that those children who are living in a single-parent family found it difficult to adapt to the school’s expectations. A teacher's assessment in 1994 comparing a group of students in intact and nonintact families, confirmed that children from disrupted families were less productive, rigid, and passive upon coping with themselves compared to their peers from the other group (Kurtz, 1994). Another group reported negative behavior changes in two-thirds of the divorces groups, as they either cannot concentrate on schoolwork, suffered from an emotional breakdown, and were in extreme need of teacher’s attention as their next best source of adult (Wallerstein & Kelly, 1980).

### 2.3.3. Expel and Suspension

A paper supported the hypothesis that the causal effect between misbehavior and poor academic performance works both ways (Myers et al., 1987). However, another study reported that if we control for instrument variables such as the principals’ preference toward discipline, and selection effects, the negative correlation between the chance of suspension and academic performance disappears (Canon, 2012).

Having investigated the rate of suspension and expel among students of disrupted families Manning & Lamb (2003) told that adolescents in cohabiting stepfather families are exposed to 122% higher odds of being suspended from school coupled with a greater chance of delinquency and misbehavior. Concerning the difference whether married or cohabiting stepfamily has larger downward effects on children on a bivariate, teenagers living in the married structure prone to have significantly less chance of being suspended or expelled from school than teens of cohabiting structure. Controlling for SES and parenting variables (marital relationship, and monitoring) the odds of suspension are the same for remarried and cohabiting types. Finally, the effect of stepfather families (either married or cohabiting) does not differ from that of their respective group in single-mother families, independently from the controlling variables (Manning & Lamb, 2003).

### 2.3.4. Grade Repeat and Dropping Out

It has been confirmed by numerous previous studies, which accounted and controlled for each kind of family background, that the structural change as early as the first grade increases the likelihood of dropping out during high school, while those students who are part of a single-parent or a female-headed household are two to three times more likely to drop out compared to the two-parent families, especially minorities. The direction of a causal effect between the breakdown in the family and the financial deprivation was considered so that the structure change causes child poverty and a possible dropout or termination of future studies (Pong & Ju, 2000).

A study between 1988 and 1992 using the data of the National Education Longitudinal Study (NELS) confirmed that the disruption of two-parent families increases the odds of dropping out of high school. In fact, for those children who have experienced a change to a single-parent, stepfamily, the risk of being a dropout in the future is two times higher. The most affected group was mother-only families, as they were strongly associated with income decrease or entry into poverty, while the likelihood of dropout was the highest (Downey, 1994; Pong & Ju, 2000).

As expected, if we control for family income change and initial poverty status, it eliminated the effect of changing to a single-parent family. Regarding the change to the stepparent family, the effect appears to be unrelated to the income change, which can be explained by the additional income of the stepparent. Students who were below the $15.000 per year threshold income, were 1.8 times more likely to drop out of school, and the risk is even higher if they experienced the drop in income during the period when the study was conducted. Thereby, both variables, the family structure, and financial stability individually and jointly contribute to the academic performance and chances of dropping out, even in two-parent families (Pong & Ju, 2000).

Undoubtedly, income plays a huge part in fostering a child’s education. Lower incomes can result in the lack of educational resources and eventually a grade repeat or in worse case scenarios a dropout. It has been shown that a percentage rase in income raises the odds of not repeating a grade, hence we can also interpret it reversely which implies that a decrease in income increases the chance of grade repeat, which is exactly what happens when a parent leaves the household. Interestingly, parental involvement does not explain the rate of grade repeat for children, which may since sometimes fewer but better consultations with parents pay off and it is difficult to measure the quality of these interactions (Kim, 2004).

### 2.3.5. Educational Attainment

Not only the family structure affects the grades in school, but also has an impact on the educational expectation. Controlling for other covariates (e.g. race, gender, parental education, socioeconomic status), children in single-parent families are less likely to finish high school or attend tertiary education (Astone & McLanahan, 1994; Biblarz et al., 2016). There is also hard evidence on the phenomena that single-parent families (especially mother-only families) tend to be less advantageous financially which is reflected in the income difference between in high school graduation between intact and nonintact families (30 to 50%), while can explain the phenomenon of educational attainment. (Astone & McLanahan, 1991).

Previous studies have shown that the educational attainment for men who grew up in a male-headed single-parent family was lower than that of men who grew up alongside their single-mother. The former group completed a year less schooling averagely compared to their cohort, while the latter group was had a disadvantage of half-year below average (Featherman & Hauser, 1978). It seems that the main problem of single-mother families is not the loss of a male partner but rather the loss of income produced by the father. In contrast, the opposite theory may be applied to the issue of single-father families, where they often miss a helping hand in coping with the children and cannot offer the same emotional support as mothers.

## 2.4. Control Variables

### 2.4.1. Age Differences

Let us turn to the discussion of control variables that affect student and school characteristics. The preschool years (aged 0 – 5.5) seems to be the most sensitive period for the children as in the case of white males and black females who spent their whole preschool period in single-parent family types, the time spent in education was shortened by half to one year, controlling for income. Such an effect for those who were in elementary or high school in the post-separation period was not significant for any race-sex groups (Krein & Beller, 1988).

Entering the school-aged group, the separation seems especially difficult because, at the age of six-to-eight, the children feel that they are rejected by one of their parents, which reflects their academic performance and diminishing well-being. At this age group, Teyber (2001) also noticed a change in the primary feeling as it shifted from sadness to anger for the nine-to-twelve-year-olds. Lastly, he argued that in the group of adolescents, the responses varied greatly, but they were generally less affected, considering they were starting to become independent and loosen their bond with the family. On the other hand, some adolescents become depressed and cannot concentrate on themselves, thereby lost their peers, ambitions, and educational aspirations (Teyber, 2001).

Regarding the time horizon of the separation, Amato (1987) claimed that negatives responses for instance general distress, sadness, confusion, etc. were more frequent if the separation occurred within five years compared to when it was six or more years ago. He also noticed that neutral feelings were more common in the age of adolescents, which may due to the better understanding of social relationships which makes them aware of the parental conflict, and can prepare themselves for the separation (Amato, 1987b). Each additional year spent in a single-parent family is responsible for a one-tenth reduction in the years of schooling for white males (significant at 1%). Those who spent an average duration in such family type (5.1 years) tend to have 0.5 fewer years of education than those in two-parent families, whereas those who spent their pre, elementary, and high school years in single-parent families (ages 0-18) had 1.7 years less education (Krein & Beller, 1988).

### 2.4.2. Gender Differences

A convincing majority of papers have shown that in mother-only families, boys are suffering more as they exhibit anti-social, aggressive behavior while having conflicts occasionally with their parents, teacher, and peers. Hargreaves (1992) claimed that boys of elementary school age who witnessed a divorce are more sensitive to the separation, while girls have a delayed reaction during their adolescence (Hargreaves, 1992).

Teyber (2001) found that in the preschool age group, both genders have experienced shock and long-term sadness. Boys in single-mother families have more adverse adjustment problems in the long-term compared to girls in single-mother families as younger ones seek more attention, and older boys are more aggressive and disobedient (Teyber, 2001). Stressful times among girls are rather associated with late adolescence and young adulthood, which can be attributed to the phenomena that when they start to arrive at the stage where they face marriage and love, they recall the memories of the unsuccessful marriage. (Wallerstein & Corbin, 1989).

The difference between the two genders can be explained by the interruption of the father-son connection, while the boys are more exposed to marital conflict. It is often assumed that parents are more involved in disciplining their same-sex child, thus when the father leaves the household, there are fewer disciplinary actions, and less authority taken towards the boy, especially if the father withdraws his interest from their children (e.g., stops paying for the childcare and visits less often). Since in the large share of cases (nearly 90%), the single-parent family is mother-only, elementary school-age boys are more likely to be exposed to the family transition and adjust slower than girls (Hargreaves, 1992).

### 2.4.3. Ethnic and Racial Differences

The effect of family structure varies across different racial backgrounds. Most of the literature believes that blacks or any other minorities are more likely to live in a single-parent family compared to their white peers. Krein & Beller (1988) provided support for such a hypothesis as black men and women were 2.5 times more likely to live in a single-parent family due to a divorce or parental death. Furthermore, they reported that blacks spend more time in single-parent structures on average, however, the negative impact is more substantial for whites especially for those who spent a greater duration in a single-parent structure (Cross, 2020; Krein & Beller, 1988).

A recent paper by Cross (2020) claimed that the more time spent in a single-parent or stepparent family the less likely the children will finish high school on time and enroll in college. Mapping such phenomena to the ethnic groups, minorities were less likely to complete secondary education on time compared to whites. The joint effect of the single-mother and race on completion on-time yielded a positive and highly significant coefficient for the black children. The difference between the Hispanic and white groups was not statistically feasible. Including the socioeconomic stressors (e.g., family income, parents’ education and age at childbirth) and family embeddedness (co-residence with external family and practical, emotional support) the effect of the interaction term diminished and happens to be no longer significant, therefore the explain most of the variation (22 to 48% and 15 to 20% respectively) between blacks and whites when it comes to high-school completion. In terms of economic resources, the average family wealth of whites is 1.5 times higher than that of black and Hispanic children. Also, whites were 1.5 times more likely to report a better neighborhood and community around them which can also foster their educational aspirations (Cross, 2020).

### 2.4.4. Birthweight and Health Status

Birthweight and adverse health status are both mild determinants of academic performance. Intuitively the birthweight should reflect the available resources of the household to support child development. Countries with a higher rate of people who saw daylight with under 2,500 grams tend to have lower GDP per capita, vice versa. Behrman & Rosenzweig, (2004) showed that indeed there is a significantly positive link between fetal growth and schooling as well as physical attributes, and earnings (Behrman & Rosenzweig, 2004; Case et al., 2005). Other evidence concluded that children with reported poor health states are more likely to be exposed to school problems and lower levels of human capital accumulation (in schooling and absolute income) both in the long and short run, which is also interrelated with the family background. The health status was captured by a categorical subjective evaluation on a scale of 1 (Excellent) to 5 (Poor), and found to be eroded faster in households with lower SES, which resulted in a downward spiral in terms of average income and years of schooling (Case et al., 2005).

### 2.4.5. School Quality

Although there is a free school choice in Hungary, households tend to choose schools on a basis of SES and expected outcomes. Therefore, we can control for the types and funders of the school along with the regional differences to helps us signal the abilities of the students. Cappellari (2004) showed that there is a relationship between the resources and high school choice among Italian youth. As far as performances are concerned, private schools are less effective than public ones in terms of educational attainment and future employment. Controlling for household income, wealth, and general health status, Newhouse & Beegle (2011) also argued that students attending public junior secondary schools have statistically higher standardized test scores upon completion compared to their peers in private school (Newhouse & Beegle, 2011). Besides, students in charter schools prone to be more proficient in both reading and math when matched to regular public school students (Hoxby, 2004). On the contrary, Lubienski & Lubienski (2006) stated that demographical differences can be held accountable for differences in math scores rather than school sectors.

# Chapter 3. - Data Source and Methods

The data used for my empirical research will be an extract from the repeated cross-sectional data from the Hungarian Life Course Survey (HLCS) of TÁRKI conducted in Hungary. It consists of the inputs of a detailed cross-sectional survey from 2006 to 2009 addressing a cohort of children starting secondary school in 2006 as 9th graders. The initial sample size was 10,022 in 2006 which decreased to 8,333 due to sample attrition. The sample is representative of the Hungarian student population, with oversampling the lower end of the social status distribution (more on the sampling see (Kertesi & Kézdi, 2011)). Approximately, 1,500 questions formed according to the National Longitudinal Surveys of Youth in the United States (NLSY79), were asked from children and parents of different households and regions. The questionnaire covers almost every corner of their socioeconomic background, past and current happenings in their life. Besides, it touches upon the cognitive and emotional skills of the children. In the following subsections, I will provide an elaborate description of the initial data recorded in 2006 and discuss the difference in the point estimates of intact and non-intact families, then I will turn to the empirical approach towards the estimation which will show us whether there is a causal relationship between the parental separation and the deterioration of academic performance.

## 3.1. Frequencies and Cross Tabulations

Starting with the time-invariant characteristics of the individuals, 53% of the adolescents were males and 36% were females (see Appendix 3.1.). Similar to Kertesi & Kézdi (2011), we will consider as Roma all students who have at least one biological parent identifying themselves as Roma either primarily or secondarily (Kertesi & Kézdi, 2011). Following such methodology, we can state that 7.56% of the children had Roma-origin. (see Appendix 3.2.). According to Kertesi & Kézdi, (2011), the difference in test scores between Hungarian and Roma children is comparable with the score gap between whites and blacks in the United States (Kertesi & Kézdi, 2011), therefore, it can serve as a feasible covariate, however, due to the lack of data we cannot explicitly confirm nor can deny the hypotheses set by other scholars.

Regarding our crucial variable, family structure, nearly two-thirds of the respondents in 2006 live in two-parent intact families while one-third of the children are part of a household that was already exposed to family disruption. Breaking down the non-intact family types, it is consistent with the preliminary literature that the custodial right is often assigned to mothers rather than fathers after the separation as the ratio of single-mother families is more than 9 times larger than that of the single-father families (6.77% versus 57.26%) (see Appendix 3.3.).

We can also investigate further family types such as different types of two and single-parent foster families, or households where the child lives without any kind of parent, but due to the small sample size, it is unlikely to conclude meaningful results. Therefore, my estimation results will only consider the parameter estimations on the pairwise difference of intact and non-intact families, where the upper bound interpretation of the “non-intact” families will be all kinds of non-two biological parent families, and the lower bound is where the children are only separated from one biological parent (conservative approach). As far as descriptive statistics are concerned, we will stick to the upper bound notion. In terms of separation type, divorce is responsible for the lion’s share of separations as it happened in 91.39% of the non-intact families (see Appendix 3.3.). Besides, parental death and other separation types also lacked in sample sizes, thus we cannot observe the main differences in the effects of separation types.

## 3.2. Central Tendencies

The average age of children at the maternal separation is 7.6 years while for paternal separation it is 6.8 years. Therefore, the separation often happens in the preschool age (see Appendix 3.5.). Considering the main determinants of academic performance, the distribution of the GPA is somewhat left-skewed, and the mean is 3.7 (at a scaling up to 5). Across the two main family types intact and non-intact, the latter’s grades and test scores were also lower which further suggests a true negative shock. Consequently, on average, children experiencing a breakdown of the family tend to achieve at lower levels, although, by a low margin (see Appendix 3.6.).

Further findings in the patterns of our data support the vast body of literature. Firstly, the mean propensity to consume (how much is needed for a month) in non-intact families is three-fourth of the propensity associated with two-parent families, which also backed by the fact that the mean net income of the mothers is considerably lower than that of the fathers (see Appendix 3.7.). Statistically speaking, when the father leaves the household in most separation cases, the average family income, and the propensity to consume decreases, which can explain the variations in academic performance.

To measure the home environment of the student, the cognitive stimulation, and the emotional support of the children, a “HOME” score was calculated for the individuals based on a shortened questionnaire of the items developed by Bradley et al. (2000), and identical to the one used in NLSY79 (Bradley et al., 2000; *NLSY79 Child and Young Adult Data User’s Guide 2002 Appendix A.*, 2004). The index incorporates the extent of parental care and involvement in the family, as well as other factors that are likely to describe the environment and encouragement within the household (see Appendix 3.8.). Mapping the average home scores across the non-intact and intact family structures, children in disrupted families tend to have lower cognitive and emotional skills than their peers in intact families (see Appendix 3.8.).

At the extremes of poor academic behavior, let us consider the repeat of school grades and the expel from school. Although the sample of those who repeated a school grade throughout the primary school is low, the proportions suggest that the likelihood of having repeated a grade (one or more times) is larger for children of non-intact families (see Appendix 3.9.). As regards the children expelled from school and dropout (leaving school due to low grades), the lack of variability does not allow us to estimate with adequate statistical significance (see Appendix 3.10., and Appendix 3.11.).

To observe the parental involvement of disrupted families, let us construct an aggregate variable. The parental school engagement score considers the parent’s frequency of attending and involving themselves in school-related matters. The higher the scores, the more frequently they involve themselves. As expected, non-intact families have lower scores, thus their parents contribute less to schooling (see Appendix 3.12.). The calculated scores were multiplied with a negative of one to reverse the order of interpretation.

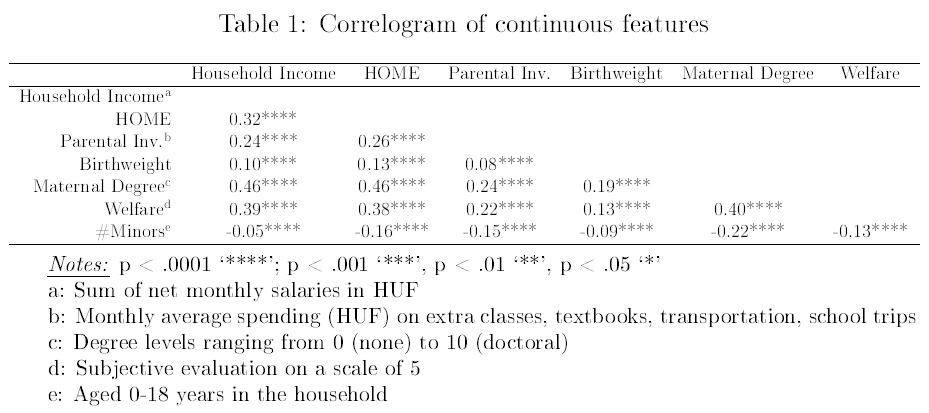
As it was mentioned in the previous sections, economic resources involve parental investments that foster education. To reflect its magnitude, I considered the monthly spending on textbooks, equipment (e.g., pens, notebooks, etc.), transportation, extra classes, and school-organized trips. In line with our preliminary expectations, intact families tend to invest more on average than disrupted families (see Appendix 3.13.).

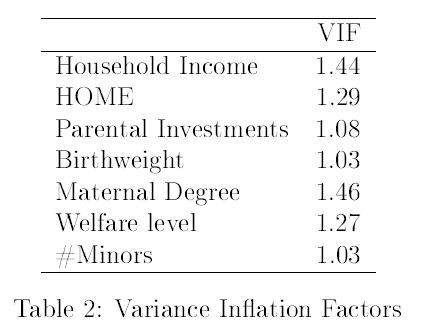
At last, I considered the residential mobility across family types and despite the small sample, the data collected in 2006 shows similar patterns as a large share of children who changed school due to moving (which was likely to be inferred by the parental separation) was part of non-intact families. Interestingly, such share increases by the number of school changes, which suggests that disrupted families may move on multiple occasions and more mobile than their counterparties (see Appendix 3.13).

## 3.3. Feature selection

To take the next step in building up our model, let us perform a feature selection by building up pairwise correlations between the underlying control variables. One of the main control variables is undoubtedly the household income in log terms. Under such a category, we can include the maternal degree level which is likely to correlate with that of the paternal as partners often find each other through homogamy. Also, income variables correlate with the school-related parental investments (e.g., extra classes, textbooks, etc.), while the HOME score is expected to move along with the previously listed measures of SES, thus, they need to be included in the picture. The last two variables which I found to be important were the reported subjective welfare of the household (*“How well does your family live compared to other families?”*), the birthweight of the student, and the number of minors within the household. Knowing that we should consider dropping regressors from the highly correlated pairs to attain the “best linear unbiased estimator” (BLUE). Observing the coefficients between the SES and HOME variables, there is a highly significant positive relationship between them to an extent (all lower than 0.5) that it should not be able to cause multicollinearity. Other recognizable patterns are the negative relationships between the income controls and the number of siblings, which is intuitive since parents are likely to partition their investments across siblings. At first glance, there is no such variable that should be dropped, hence we can reject multicollinearity and assume that there is only a negligible bias between the effects of the variables (see Table 1.).

Put differently, strong exogeneity can be achieved by including the enumerated continuous variables. In the next section, I will first discuss the model selection by implementing a forward selection strategy, which then will arrive at the ultimate model with the full set of both categorical and continuous variables which may account for the fundamental assumptions of the OLS.



To be more confident in including these variables at the right-hand side of our future specifications, we can observe the Variance Inflation Factor (VIF)) of the reduced data frame containing the continuous variables. According to the rule of thumb, a variable should be omitted from the equation if its VIF value exceeds the cutoff value of 5.0, which in this case is not applicable, therefore they pass both stages of the baseline screening.

## 3.4. Methods

### 3.4.1. Value-Added Model (VAM)

Value-Added Models (VAMs) are one of the most used tools when it comes to understanding educational production. It first saw daylight when assessing student outcomes associated with the No Child Left Behind Act (NCLB) of 2011, and became widely adopted after the paper of Sanders, Saxton, & Horn (1997) on Tennessee jurisdictions (Wright et al., 1997). In our case, it tells us how much value a separation added to the lagged-GPA of a student in 2006 as of 2009. The fundamental purpose is that lagged school grades (benchmark) in the regression equation incorporate most of the student and school covariates (not under the control of parents) between the control and treatment groups. Along with the lagged performance, home environment, SES, PSI, residential mobility, and the approximate time of separation variables are also included. Such a methodology allows us to represent the static difference between family types in a one-shot estimation and confirm that negative patterns are recognizable after the family transition. Given that our results will be significant and in line with the existing literature, should we proceed to the dynamic approach described in the next subsection where we account for the length of exposure. As a separation indicator category, we will distinguish three groups: children whose biological parents (1) were separated before 2006, (2) between 2006 and 2009, (3) or never experienced a disruption.

I will construct seven specifications where I build upon the unconditional model based on the forward selection principle. In the first phase, I will estimate the baseline model without any discussed covariates. Next, I will add regressors one by one starting with the set of exogeneous effects (student characteristics, SES) which are independently responsible for the outcome variable and not induced by parental separation, then continue with mitigators (home environment, school characteristics, PSI, residential mobility) which are often referred to as the channels of family disruption effect. The Directed Acyclical Graph (DAG) (Figure 1.) represents the effect flows of each variable and illustrates my hypothesis when building the models.

The full (seventh) specification of our linear VAM is demonstrated as follows,

|  |  |  |
| --- | --- | --- |
|  |  | (1) |

where we regress the reported GPA in 2009 onto the (1) constant, (2) the prior grade reported in 2006, (3) the set of student characteristics in 2009 consisting variables of Roma-origin, sex, age, birth weight, and reported health status, (4) the determinants of SES (log household income, maternal degree, reported welfare), (5) home environment (HOME index, #minors in the household), (6) the school characteristics, for instance, the maintainer (charter, church, foundation, other), and type (vocational or equivalent, high-school, other) of school. To continue, I will also add (7) the parameters representing the PSI (school engagement, log monthly investments), and the (8) number of school changes due to movement before 2006 as a covariate of residential mobility. Finally, define the ternary separation indicator for each student . and let denote the idiosyncratic error term of the estimation. For the full description of each variable, please refer to the appropriate section of the Appendix (*Description of variables*). In a way, the design above isolates the separation effect contribution from all the outside factors affecting the performance. We expect negative coefficient signs for separation categories describing separation in decreasing trend close to Diagram

Description automatically generatedzero as we include more and more sets of covariates.

Figure 1. Directed Acyclical Graph (DAG) of variables used in this study.

It is crucial that whenever we include a covariate, there may be observations that have missing values, therefore there is a trade-off between uniformity of the groups and efficiency, and the confidence intervals are further stretched out. To assess the accuracy of our models, I will compare the penalization statistics (Adjusted RSquared, Akaike Information Criterion (AIC), Baynesian Information Criterion (BIC)) and pick our best estimator. One would argue to include school-specific IDs as a proxy to the school quality fixed effect, but unfortunately there was too much variation in the schools which was penalized by the BIC (tripled the score) and resulted in insignificant findings.

### 3.4.2. Staggered Difference in Differences

To evaluate the dynamic effect of family disruption on academic performance, let us implement the method of Staggered Difference in Differences (DID) with multiple periods developed recently by Callaway & Sant’Anna (2019). I will fully rely on their approach since they provided more flexible and relatively robust staggered adoption compared to previous papers. Furthermore, it is convenient to apply the model to hands-on data since Callaway developed an open source for the R software for the purpose. In what follows, I will introduce the statistical approach and the proposals for treating the known issues with DID designs and offer a mathematical representation of the main advantages of the augmented dynamic structures.

#### 3.4.2.1. Known Issues of the Standard DID

The main distinct assumption of the Difference in differences method is the so-called Parallel Trend Assumption (PTA), which requires that the measurement variables be either time-invariant group attributes or time-varying factors that are group invariant. Jointly, these imply that the time series of academic outcomes in the group of intact and non-intact families should differ by a fixed amount in every period and exhibit a common set of period-specific changes. Although, in a simple two-group two-period design the assumption is not testable, there are graphical methods that allow us to show that the trends across groups are parallel to each other (Wing et al., 2018).

Bertrand et al (2004) argued that there are also other concerns related to standard errors of the estimations due to heterogeneity. A significant problem that may arise and often ignore by researchers is the notion of serial correlation, which overestimates the t-statistics and significance levels in the DID, but in our case, it is not likely to occur since we have a relatively short time-series dimension (). However, the short time series length of the classical misspecifies the process may lead to inconsistent standard error terms. If it would be the case, either a placebo parental separation by Monte-Carlo simulation (or bootstrapping) or an autocorrelation ( structure of the error terms should correct the phenomena (Bertrand et al., 2004).

What is often criticized in the “static” TWFE model is that the coefficient of DID (interaction of time and year) does not reflect the causal parameter of interest when there is a variation in timing and the effects are dynamic. Rather, it represents the weighted average of dynamic effects, allowing for negative weights in the long-run, therefore it creates bias and “overweight” the short-run effects. Also, it heavily relies on the homogeneity of treatment effects (Callaway & Sant’Anna, 2018).

Another important principle is connected to strict exogeneity which states that the DID design should aim to difference out unmeasured confounders, by eliminating group-specific and time-invariant biases. That is, the time factor must be statistically independent of the outcome distribution across the lagged periods, conditional on the group- and time-fixed effects. Other assumptions are not specifically connected to the DID, but rather the linear model of OLS which is constituted under the Gauss-Markov assumptions (Wing et al., 2018).

Finally, scholars also mention that the distribution of covariates should be very similar in both family structure groups. The incentive is to show that the two groups were comparable before the parental separation happened. In the case of DID design, it is only necessary that the differences between the two groups are stable over time and that the treatment is not responsible for the changes in the distribution of covariates. To test such a theory, one can replace the outcome variable with the covariate and fit the baseline DID regression model. If the DID coefficient is near to zero or not significant, it suggests that there are no compositional changes (Wing et al., 2018).

#### 3.4.2.2. Specification

Callaway & Sant’Anna’s (2019) procedure considers multiple time periods with variations in treatment timing, while it accounts for the PTA conditioned on observed covariates which makes it unique compared to other models. The strategy can be divided into two stages, where the first establishes the asymptotic properties of the treatment estimator, and the second implements a bootstrap procedure to carry out asymptotically valid simultaneous inference. Bootstrapping allows for simultaneous confidence intervals instead of the traditional pointwise estimation and lowers the likelihood of heterogeneity. To pre-test, the validity of the PTA, a fully data-driven semiparametric t-test procedure offers credible solutions, which is based on the weighted integrated conditional moments (ICM) approach introduced by Bierens & Ploberger (2016) (Callaway & Sant’Anna, 2018).

Instead of the coefficient in the TWFE regressions, Callaway & Sant’Anna (2018) proposed the group-time average treatment effects (), which accounts for the time of first treatment and the evolution of dynamic effects, thus mitigates heterogeneity and makes it easier to interpret the causal effects. Callaway & Sant’Anna (2018) identified three possible alterations when aggregating the s, which are the (i) selective treatment timing, (ii) the dynamic treatment effect, and the (iii) calendar time effect. In this study, I will only consider the dynamic effects which assume that the effect of parental separation may depend on the length of exposure. Therefore, by aggregating across groups and periods, we will have one robust parameter estimate of the effect of parental separation on academic performance (Callaway & Sant’Anna, 2018).

Let us begin with the composition of groups. Our control group incorporates the children with intact families, and the group of children in non-intact families (by definition) provides the treatment group. Since the DID method observes differences in the group-time averages, repeated cross-sectional data will suit the purpose and we do not need to treat the issue of unbalanced panels. We will narrow down our observation period for the years between 2006 and 2009 ( when students were likely to attend secondary education, having 2006 the pre-treatment year. Observations that were already treated in 2006 are dropped from the sample. As it was earlier emphasized, parental separation can occur at different points of time between the span of four years and the recorded year of treatment will be fed into a variable, therefore we can accurately determine the year of separation.

In what follows, I will briefly describe the key points of the approach, since the proofs and remarks of the theorems are beyond the scope of this essay. Should you deeply interested in the details, please refer to Callaway & Sant’Anna (2018). In our case, there are periods in total between 2006 and 2009, where in 2006 () no one is yet treated. Let be the binary variable of treatment which is equal to one if the children live in a non-intact household and zero otherwise. will also be a binary variable that is equal to one if the first separation occurs in the period of , given . Also, is a binary variable that takes a value of one if the individual is in the control group (never experienced a family disruption within the four years). The dummy of and fully complements the population, as for each individual exactly one of the variables is equal to one. Define the generalized propensity score constructed by the authors as , where is the estimated probability that a student is treated conditional on covariates of , group . Lastly, denote and as the potential outcomes at time t with and without treatment, respectively. Knowing that, the observed outcome can be described as . Given that we are not able to observe both outcomes for each individual, the group-time average treatment effect () offers a better solution, which has the following functional form (Callaway & Sant’Anna, 2018):

|  |  |  |
| --- | --- | --- |
|  |  | (1) |

Such implementation of the treatment effect offers flexibility to address different dimensions of heterogeneity by the time of adoption. Subsequently, it allows us to interpret the effect of parental separation over time and compare the short and long-run effects.

#### 3.4.2.3. Assumptions

Before discussing the details of the nonparametric identification strategy for the , let us summarize the main assumptions indicated in the paper of Callaway & Sant’Anna (2018).

**Assumption 1 (Sampling for Repeated Cross-sections):** *Conditional of , the data are independent and identically distributed (iid.) from the distribution of , for all . That is our sample consists of random draws from the mixture distribution* (Callaway & Sant’anna, 2019)*.*

**Assumption 2 (Augmented Conditional Parallel Trends):** *The conditional PTA should hold in multiple periods for each group, so the average outcomes of the groups are equal in absence of the treatment. Formally, for all , such that* (Callaway & Sant’Anna, 2018)*,*

|  |  |  |
| --- | --- | --- |
|  |  | (2) |

**Assumption 3 (Irreversibility of Treatment*):*** *Once an individual is treated (separated from one parent), he/she will remain so in the next period* (Callaway & Sant’Anna, 2018)*. That is, for , implies that*

**Assumption 4 (Overlap):** *Positive share of the population starts to be treated in period , and given covariates of there is also a positive probability that some are not treated* (Callaway & Sant’Anna, 2018)*. Formally, for all and for some*

#### 3.4.2.4. Group-time Average Treatment Effect

Under the assumptions above and for , the group-time average treatment effects () for group and in period , condition on covariates can be nonparametrically identified by the following theorem (Callaway & Sant’Anna, 2018)

|  |  |  |
| --- | --- | --- |
|  |  | (3) |

Translating the equation into reasoning, a simple weighted average of “long differences” in the outcome yields us the group-time average treatment effect (). To construct the weights, one should take the observations of the control group (intact families) and group (exposed to parental separation in period ), omit other groups, then weigh up observations from the control that have similar characteristics to those in group . Similarly, observations that have characteristics rarely found in group should have lower weights. Such adjustment guarantees that the covariates are balanced in both groups, therefore outliers do not induce heterogeneity. The estimation has an unconditional case when , that is that we do not include any control variables into our model, but it is beyond the scope of this paper to elaborate its methodology and can be easily derived by substituting the appropriate value. The final form of the denoted as follows, which is much simpler than the conditional case in (3) (Callaway & Sant’Anna, 2018).

|  |  |  |
| --- | --- | --- |
|  |  | (4) |

Thes can be aggregated to obtain a single measurement of parental separation for the sake of causal inference. Aggregating the dynamic effect increases statistical power and reduces estimation uncertainty. One way to combine the effects across and is to take the sum of weighted averages, where more weight put on the with larger group sizes. Unlike in the case of TWFE where the can consist of negative weight, such aggregation rules out the bias in the sing of the coefficient (Callaway & Sant’Anna, 2018).

|  |  |  |
| --- | --- | --- |
|  |  | (5) |

where which ensures that the sum of weights is one.

#### 3.4.2.5. Dynamic Treatment Effects

In our case, we assumed that the effect of parental separation is dynamic, put differently, it depends on the length of exposure. We hypothesize that the initial effect of the separation in period is the strongest, and afterward, the children manage to readjust slowly. Let us denote as the aggregation of across the length of exposure to family disruption.

|  |  |  |
| --- | --- | --- |
|  |  | (6) |

The equation above reflects the average effect of treatment for children been in non-intact families for exactly periods. If we aggregate the possible values of , it yields the dynamic effect measurement (Callaway & Sant’Anna, 2018)

|  |  |  |
| --- | --- | --- |
|  |  | (7) |

#### 3.4.2.6. Bootstrapping

To solve the problem of endogeneity of intervention, serial correlation, and to set up the estimation of the propensity score given the covariates , I will also implement the augmented bootstrap procedure introduced by Callaway & Sant’Anna (2018). It is called augmented because it leverages the asymptotic linear representations and, in each iteration, there are always observations from each group (2007-2009), unlike in traditional bootstrap where there may be missing groups in each draw. A further advantage of the simple multiplier bootstrap is that it uses simultaneous confidence bands which catches all the with a probability of at least , where is the significance level. The parametric implementation of bootstrap extends beyond the paper, therefore it will not be discussed further (Callaway & Sant’Anna, 2018).

#### 3.4.2.7. Pre-testing the Conditional PTA

Previously we introduced a powerful design that can nonparametrically identify asymptotically valid inference about the causal effect of parental separation on academic performance, using conditional (and unconditional) multi-period dynamic DID models. Although, all the above are heavily dependent on the Conditional PTA, which is fundamentally untestable. However, Callaway & Sant’Anna (2018) came up with a stronger version which assumes that Conditional PTA holds for all periods (including treatment periods), not only for (pre-treatment periods), thus it can detect a broader set of violations of the more rigid condition. Also, they argued that the “curse of dimensionality” does not allow for feasible testing, therefore we should build our testing (Wald-test) on the ICM approach which was first used by the authors (*H0: The conditional PTA holds for all periods*) (Callaway & Sant’Anna, 2018).

#### 3.4.2.8. Forward selection

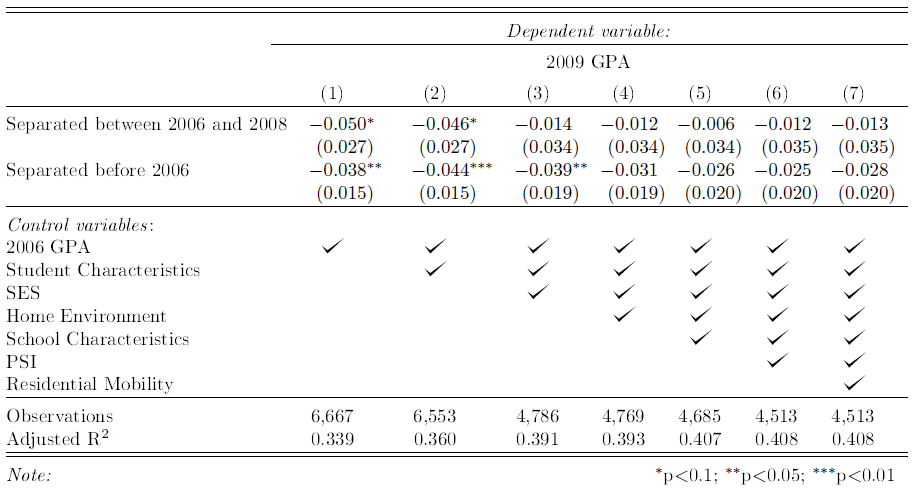
I will follow a forward selection process by incrementing sets of control variables which should help to explain a reasonable amount of variance in the outcome variable. Our baseline model will be an equation where only the family structure and the year of separation are on the right-hand side, with no covariates, thus we will estimate an unconditional s. Next, I will start to increase the complexity by adding the same sets of control variables (covariates) used in the VAM estimation. Firstly, I will append the exogeneous variables such as the student characteristics and SES, then consider mitigating factors of parental separation by including the home environment, school characteristics, and PSI variables. Finally, in our ultimate model, I will further extend the design with the regressor of school changes before 2006 which should affect the attitude of both teachers and students according to Astone & McLanahan (1994). All in all, the procedure of adding covariates will be identical to that of the VAM’s forward selection.

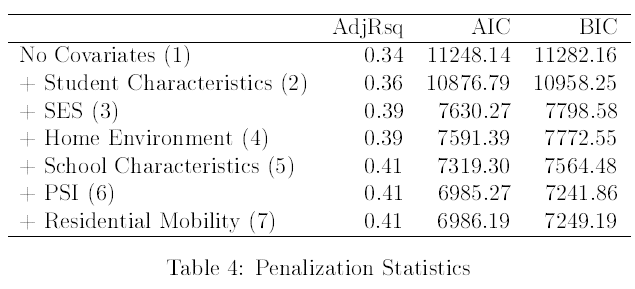
#### 3.5.2.9. Hypotheses

Before discussing the main results in the next section, let me introduce our expectations concerning the signs and magnitudes of coefficient estimates. We expect a changing group-time average treatment effect of parental separation for each group in each period depending on the time of exposure. We anticipate that the initial effect, that is the effect in the first period having exposed to family breakdown will have the largest effect with a possibility of it being a positive shock. In the long-term, it is ambiguous whether the academic performance will worsen, or children will be able to adjust to the situation. Intuitively, it can also be the case that there were conflicts in the household beforehand which made children prepare for the worst-case scenario, thus the separation will have a positive impact on their grades, as they experience relief from stress in the short run. Since we are planning to include various covariates, it can be the case that they are fully mediating the effect associated with the separation in the short run, leading to a zero-intervention impact. Thereby, let me illustrate four possible scenarios: (1) The separation has a negative effect in all the post-treatment periods, (2) it has an initial negative effect, and then the adjustments begin, so that the grades in the post-treatment periods are better (or less worse), (3) Initial positive impact, than a steady decline on the short run, (4) or a close to zero effect on both post-treatment periods.

# Chapter 4. Results

## 4.1. Value-Added Model (VAM) Estimation

In sections 4.1 and 4.2, I will discuss the results of both the VAM and the Staggered DID’s upper-lower bound estimations. Let us start with the representation of the static results which should verify the baseline belief set by many other studies. Table 3. shows us the regression output of the VAM equation, where each column is a model with additional controls. Statistically speaking, children whose family was intact in 2006, but non-intact in 2008 tend to experience a roughly .05-point drop in their GPAs (p < 0.1) if we do not condition its effect on any of the covariates other than the lagged-GPA. Regarding those who were already part of a non-intact household in 2006, should have a milder negative effect of -0.038 (p< .05) on their performance, holding all other factors at constant. As we start to include the sets of covariates, the statistical significance of the separation decreases (p > 0.1), and the magnitude shrinks to a territory floating around -0.012 for the recently separated and -0.03 for the other group. To interpret the difference between these two estimates, when the separation occurred recently, students may be under a shock and did not experience the consequences, while in the other case, they start to admit the (partial) loss of a parent which affects their performance. The group of children who have never been “treated” is dropped to avoid perfect multicollinearity. All in all, the parameter estimates confirm that there exists indeed a negative effect, although there is larger uncertainty, and the effect implies that covariates absorb most of the negative impact.

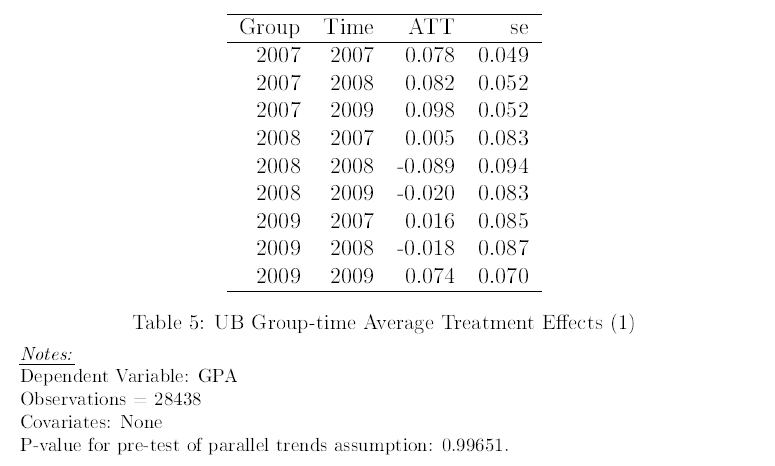
To pick our best VAM model, Table 4. consist of the main penalization statistics for each model selection. As we can see, the overall accuracy of the model increases, and the sweet spot is reached in the sixth model, having the lowest AIC and BIC scores, and calling for skepticism about the residential mobility as we have seen in the literature review. Following the same line of thought, we can proceed to the result of the Staggered DID which should give us more detail about the relationship based on the length of exposure and varying treatment periods.

## 4.2. Staggered DID Estimations

What follows are the dynamic results of the upper and lower bound estimations of parental separation on children’s GPAs through a forward model selection procedure based on our forward selection. Regarding covariates, I will mirror the explanatory sets used in the VAM estimation. Heavily relying on the dynamic approach developed by Callaway & Sant’Anna (2018), I used their R package to implement multiple Staggered DID estimations on the HLCS dataset, where the inputs required the outcome variable, the dummy for family status, the control variables, and lastly a column which indicates which year an individual was first exposed to disruption. The very last input of the model allows us to elaborate on the coefficient sizes based on the length of exposure. The estimations were carried out with clustered bootstrapped (n = 1000) standard errors at a student level, while it accounts for the autocorrelation of the data.

### 4.2.1. Upper-bound Estimates

#### 4.2.1.1. Unconditional Design

First, let us consider the result of the upper-bound estimations. In Table 5., you can see the average group-treatment effect of the parental separation given that there are no other covariates (N = 28,438). There are no outstanding groups that were substantially more affected given time, however, students of the 2007 cohort who were separated at least from one of their biological in either 2007, 2008, or 2009 performed marginally better after the separation. It is counterintuitive, since we expect negative coefficient signs, but also can be attributed to the fact that in many cases the parental conflict precedent to the disruption imposes a larger negative effect on children than the separation as they are already aware of the worst-case scenario and eventually relieved from the stress, leading to a positive separation effect. As we can see, the unconditional PTA does hold as there is no evidence against it when performing the Wald-test (p > 0.997), but the coefficients are likely to be biased as possible regressors end up in the error term. We should expect the p-value to decrease in the future as trend variances throughout the observed period show up after conditioning on controls with greater dispersion. Beginning from the next subsection, we will mitigate such issues by including sets of covariates explaining such differences and try to find the optimal trade-off between PTA and accuracy of the coefficients.

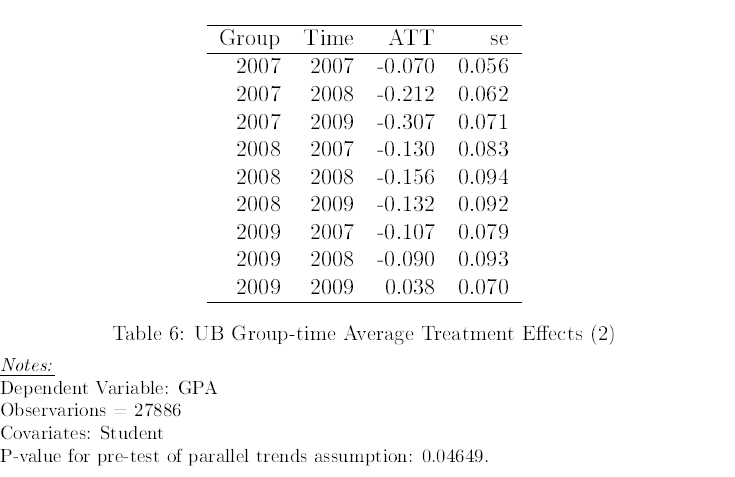
We can also plot the dynamic s and transform them into an event-study like representation where we illustrate the s on the dimension of the period and assigning zero to the time of treatment. In Figure 1., the unconditional aggregated effects through the length of exposure suggest that after the separation (t = 0), children tend to perform better on average, and in the second post-treatment year their GPA was roughly 0.1 points better. The fitted 95% confidence bands after the bootstrapping (n = 1000) include zero, thus we can also state that there is already a possibility that parental separation is irrelevant. To sum up, the dynamic effects also show a positive effect in the short run even though it is Chart, line chart

Description automatically generatedour most underspecified model.

Figure 2. Event study of parental separation’s upper-bound and unconditional impact on GPA

#### 4.2.1.2. Student Characteristics

By including the student characteristics of sex, age, Roma-origin, and health measures, there are noticeable changes in the signs and magnitude of the coefficient estimates. Nearly all the group-time coefficients have a negative sign and there are group-time pairs such as the 2007-2008 and the 2007-2009 where the effect was substantially larger than that of the other pairs. Put differently children of 2007 who were separated in 2008 or 2009, experienced an average decrease of 0.212 and 0.307 in their GPAs, respectively, ceteris paribus. In line with our expectations and the variance within the cohort, the pre-test of the PTA is significantly worse than it was in the unconditional case, meaning that we can reject PTA (p < .05) and the coefficients are weakly reliable. Thereby, the conditional results support a body of literature that parental separation imposes adverse negative effects on children’s academic grades when we account for student characteristics, but the assumptions of DID are violated (Table 6.)



Looking at the dynamic effects on the axis of exposure length, we can conclude that the negative effect associated with disruption puts a downward spiral on children as they are already performing worse in the pre-treatment period (violating the PTA) by roughly 0.1-grade point due to the constant stress level, and when the separation occurs, they fall behind with a relatively steep slope having approximately 0.3 lower average GPAs in the second post-treatment year compared to the children of intact families, holding student characteristics constant. The same hump-shaped impact as in the unconditional specification can be recognized at the separation period which can be attributed to the initial relief from pressure. It is worth mentioning that the negative effects are significantly different from zero at a 5% level in the post-treatment periods, as zero is not part of the confidence intervals (Figure 3.).

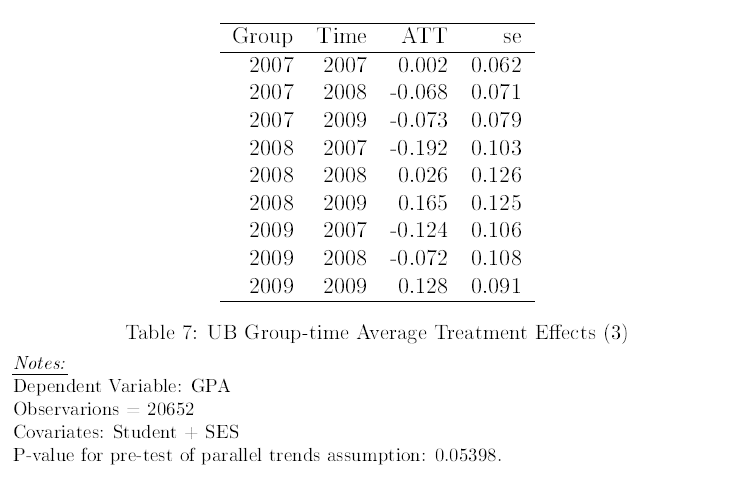
*Chart, line chart

Description automatically generated*

Figure 3. Event study of parental separation’s upper-bound impact on GPA when conditioning on Student covariates.

#### 4.2.1.5. Socioeconomic Status (SES)

Next, we should add our second exogeneous variable reflecting an individual’s SES and observe whether it can help to explain most of the negative effects as it was thoroughly emphasized in the selection of papers. I decided to add the control of the household income in logarithmic terms, the maternal degree (homogamy is assumed), and the reported welfare status of the family reported by the respondent.



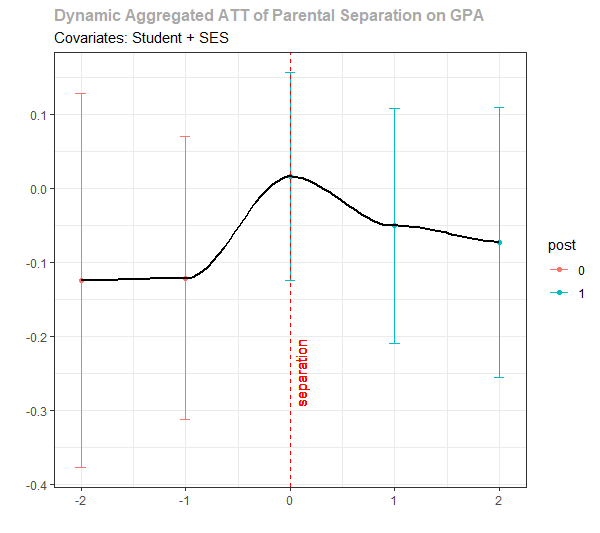
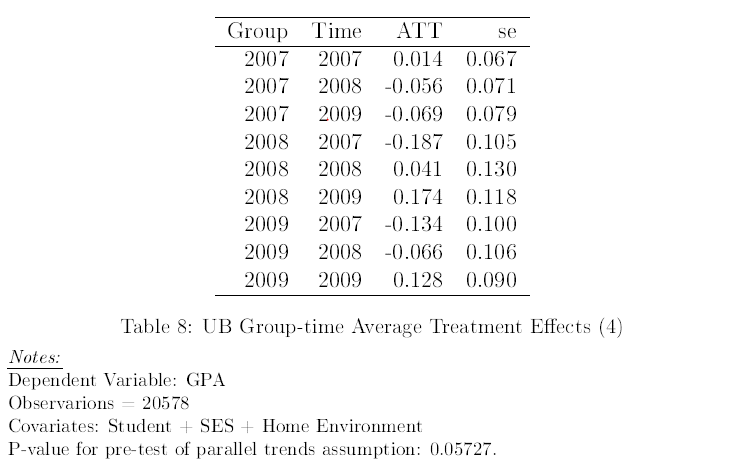
By adding the SES variables, majority of the treatment effect drops in size which is analogous to out hypothesis that it possesses the largest explanatory power (Table 7.). Regarding the PTA, SES highlights differences in the pre-treatment period and jeopardizes the assumption to hold but not to an extent that we can reject it on a 10% level (p > 0.054). Interestingly, there are larger negative effects before the separation and larger positive effects when arriving at the separation, which implies that the stress effect beforehand is even greater, resulting in a sigmoid-like representation. After that, the shock is corrected but remains in the negative half in both years after the disruption with a negative trend, suggesting that students are less affected in the short run when we also account for SES variables. Besides, all of the confidence bands covering the zero-treatment effect at 5% therefore parental separation may not be statistically meaningful at the end (Figure 4.).

Figure 4. Event study of parental separation’s upper-bound impact on GPA when conditioning on Student, and SES covariates.

#### 4.2.1.4. Home Environment

As we have seen before, the home environment plays a huge part in developing children’s cognitive and emotional capabilities. For that purpose, I will also include the HOME scores reported in 2006 and the number of minors (0-18 years of age) in the same household to capture the circumstances surrounding the individual at home. I expect that the attention and care the child receives as well as whether the adequate conditions to study is established at home does have a meaningful place in the equation and acts as a mechanical variable. Also, the number of children in the household should be correlated with the attention and investment partitioned to each child since by having lots of children, parents may not be able to satisfy the needs of everyone in the household given the time and effort constraints.

The table output below suggests that home environment covariates further mediate the negative effect of parental separation. However, the pre-test of conditional PTA is failed to be rejected at a 5% level, meaning that the fundamental assumption of DID is likely to hold and the coefficients mirror the actual values given the input variables (p > 0.05) (Table 8.).

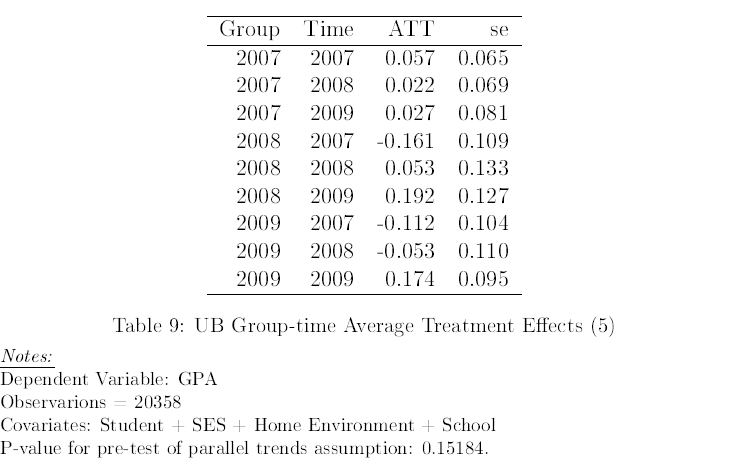
Regarding the dynamic aggregated effects, we can also observe shallower slopes in the post-treatment periods, where the decrease in average GPAs of children with disrupted families stays above 0.1 points two years after the separation. The hump shape, in this case is more expressed because the home environment somewhat measures the conflicts and strength of relationships at home (Figure 5.).

Chart, line chart

Description automatically generated

Figure 5. Event study of parental separation’s upper-bound impact on GPA when conditioning on Student, SES, and Home covariates.

#### 4.2.1.5. School Characteristics

To compensate for the weakness of GPA compared to standardized test scores, we should consider adding the school characteristics because, in each school, the expectations and the quality of the curriculum are different. Therefore, I added the variable of regional differences, the type of school, and the maintainer of the school (same categories as in the VAM (see. Section 4.1.)). The maintainer accounts for the differences in school financing which should heavily correlate with the ability to attract better teachers and offer a large variety of equipment’s e.g., electronic gadgets.

Chart, line chart

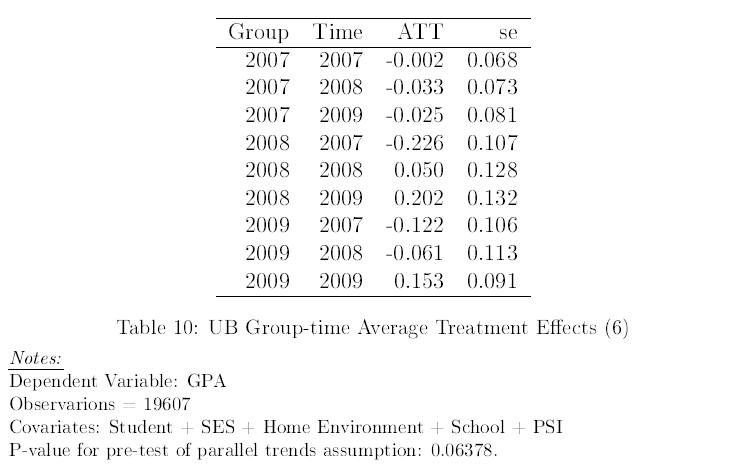
Description automatically generatedTable 9. illustrates the treatment effects when conditioning on the student, SES, home, and school controls. Such effects helped to decrease the variance across group-time pairs, while the effect shifted upwards to a positive territory which is also present when plotting the dynamic effects (Figure 6.). Regarding significance, the p-value takes up the highest value (p > 0.15), indicating no evidence even at 10% against the parallel trend assumption of the two groups, because by including such effects, it helps us to explain the main confounders of the pre-trend analysis.

Figure 6. Event study of parental separation’s upper-bound impact on GPA when conditioning on Student, SES, Home and School covariates.

#### 4.2.1.6. Parental School Involvement (PSI)

Our sixth model will increment the previous model with the determinant of PSI, meaning I will add the reported frequency of participating in parents' meetings at school and consulting with the child’s teachers along with the monthly net spending on school-related activities fostering education and well-being of the child.

There are no meaningful changes in contrast with the previous setup except the fact that the conditional PTA no longer holds at 10% level (p > 0.064) due to the difference in the level of school-related engagements and investments across control and treatment groups in the pre-treatment period (Table 10.).

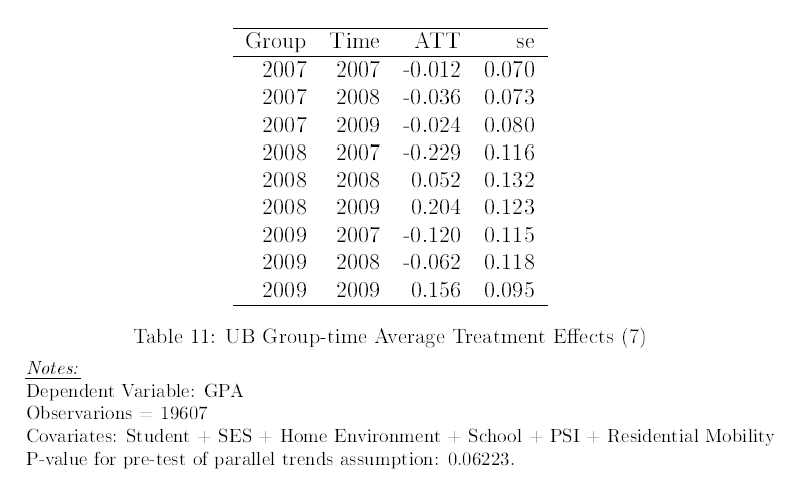


Chart, line chart

Description automatically generatedAs far as the dynamic effects are concerned, its shape is comparable to that of the SES model, but students start from marginally larger negative effects in the pre-treatment period and end up close to zero-effect after the separation with a minor negative trend, possibly because rather the conflict and stress precedent to the separation (divorce) is responsible to the negative impact, and only after a while children realize the loss. (Figure 7.).

Figure 7. Event study of parental separation’s upper-bound impact on GPA when conditioning on Student, SES, Home, School, and PSI covariates.

#### 4.2.1.7. Residential Mobility

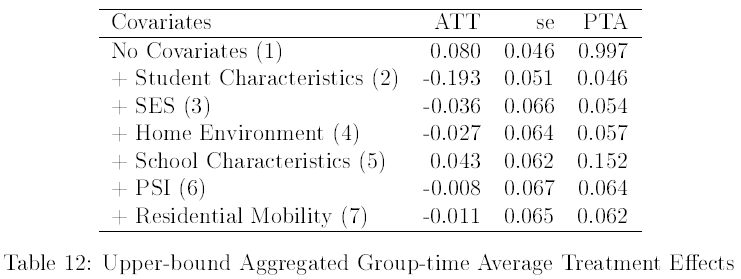
Finally, we will see how our model changes if we add the number of school changes due to moving before 2006. We expect more accurate coefficients because as relocation can be induced by a separation, thus children with high residential mobility rates can have problems at school as described in the literature review.

Chart, line chart

Description automatically generatedAccording to Table 11., and Figure 8., the average effects are nearly identical to that of the former model, which can be attributed to the lack of variance in the number of school changes. That being said, the p-value of the pre-test of the conditional PTA assumption decreased, but there is still no evidence against it at a 5% level (p > 0.062).

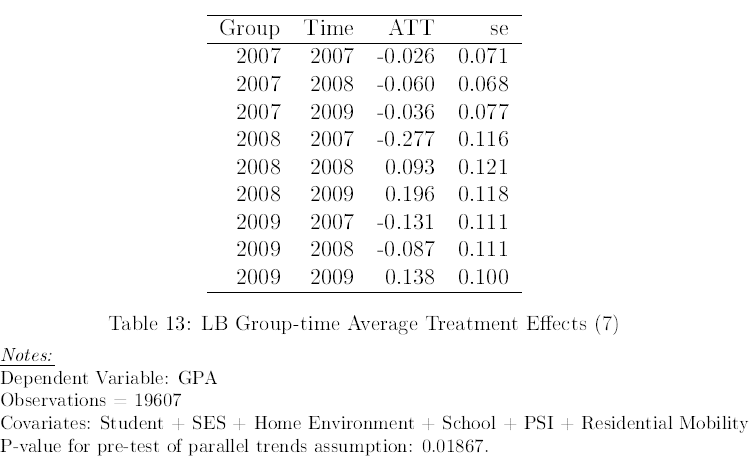
Figure 7. Event study of parental separation’s upper-bound impact on GPA when conditioning on Student, SES, Home, School, PSI, and Residential Mobility covariates.

#### 4.2.1.8. Discussion of the Upper-bound Estimates

Finally, we can aggregate the s for each model to gain one-metric estimates. For the sake of simplicity, we can compare our aggregated coefficients and summarize the findings of our forward selection. Table 12. illustrates the specifications, the aggregated s, the standard errors, and the p-values of pre-testing the (un)conditional PTA for all the seven upper-bound specifications. When there were no covariates whatsoever, separation is associated with slightly higher grades, but the effect was likely to be biased due to the lack of covariates. As I started to include different sets of covariates, the likelihood of PTA holding decreased but the more accurate coefficient estimates took up greater negative values then recovered to the level of -0.011 average treatment effect when student, home, school, SES, PSI, and residential mobility controls were held constant (7). There is clearly a tradeoff between accuracy and credibility in our models, but we can conclude that we fail to reject the PTA on a 5% level in our most complicated design, implying that parental separation does have a minor negative effect by itself on GPA, even though the included exo-and-endogeneous variables are responsible for the lion’s share of the drop in academic performance. Also, Figure 7. implies that the negative effects are marginally stronger over the years of exposure.

### 4.2.2. Lower-bound Estimates

Having estimated the negative effect of transitioning from a two-parent biological family from any kind of disrupted family where at least one original parent is missing, we can turn to the conservative approach where I will discuss the impact on the academic performance of children who were separated from only one of their biological parents. Let us invoke the fully equipped Staggered DID specification which consists of the set of student, home, school, SES, PSI, and residential mobility covariates and apply it for the restricted set of family types, rather than non-intact families in a broader sense.

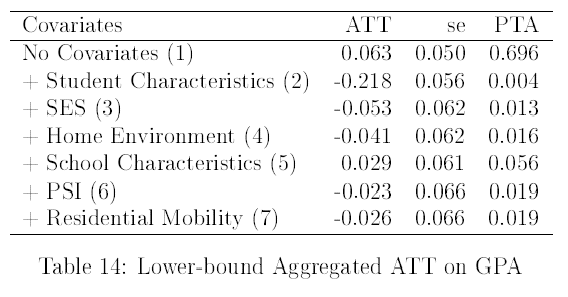
Table 13. illustrates the results of the estimation with the full set of controls. Comparing its coefficients with that of Table 11., they suggest that the negative effects are enlarged while the positive effects are milder in the conservative definition of non-intact families, therefore the effects are worse. In the latter, the conditional PTA does not hold under the 5% significance level (p > 0.019), therefore the upper-bound model is more feasible for interpretation.

Chart, line chart

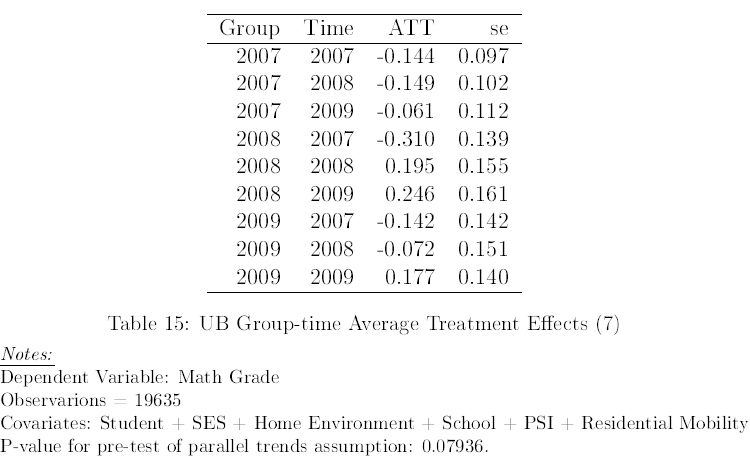
Description automatically generatedWe can also plot the event study of dynamic effects in Figure 8. and compare it to the case of extreme non-intact interpretation. The patterns are very similar to the fully equipped upper-bound estimation but at the lower-bound, the figure is shifted downwards verifying our interpretation at the group-time pair level.

Figure 8. Event study of parental separation’s lower-bound impact on GPA when conditioning on Student, SES, Home, School, PSI, and Residential Mobility covariates.

Although we omitted the unconditional and the steps of arriving at the full set DID due to efficiency, their aggregated s should be adequate to compare and conclude the differences between the definitions of non-intact families. Table 14. summarizes the comparable metrics defined in Table 12., where the same thoughts can be formed as in the group-time estimations for the full set, considering that separation affects more those children who are only separated from one of the biological parents (effect more than doubles for the full model). As far as the hypothesis testing of PTA is concerned, these results are less significant than their comparable coefficients in the upper-bound case, because, in the last two models, the identical trend in the pre-treatment period can be rejected at the 5% level.



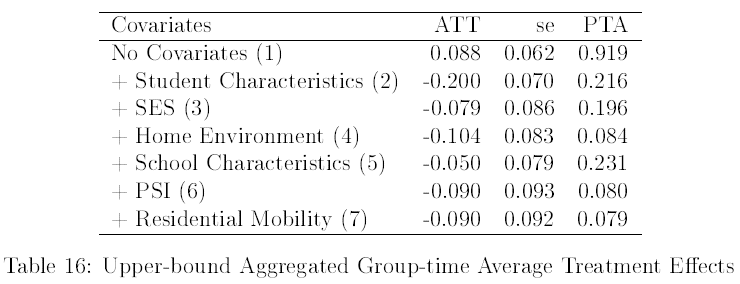
### 4.2.3. Additional Results

For additional results, let us estimate the s for the math grades of the underlying cohort. The reason why I chose math grades particularly, can be attributed to the fact that math is considered one of the most demanding subjects as it needs a lot of effort and practice to understand the logic behind it. Also, it is an objective subject so there is not much room to rely on other skills other than calculus. Not many students have an analytical mindset, therefore we expect that the grades are the most volatile among all other subjects due to their sensitivity to social status-related factors. For this purpose, I will denote non-intact families according to the upper-bound approach, due to its statistical power in terms of PTA. Table 15. elaborates on the s, from what we can derive that the aggregated s are indeed substantially more volatile than it was on the GPA as both negative and positive group-time effects on math grades are more expressed. Comparing the PTA statistics with the previous residential mobility models, we can conclude that we can achieve the more robust estimates when replacing GPA with math grades (p > 0.079), which is intuitive since GPA is an aggregated outcome variable summarizing different subjects in terms of sensitivity.

Chart, line chart

Description automatically generatedAs for the dynamic aggregated effects, the pre-treatment periods have similar negative effects on the students, although the relief from the household conflict (referred to as a hump) is hardly noticeable on the graph implying that within one year of the separation, they are still performing below par due to the absence of a parent. Also, there is a slower adjustment to the situation, and we can observe a recovery effect unlike in the case of GPA where there was a negative trend in both post-treatment periods (Figure 9.).

Figure 9. Event study of parental separation’s lower-bound impact on Math Grade when conditioning on Student, Home, School, SES, PSI, and Residential Mobility covariates.

Table 15. below summarizes the main three attributes for each specification, where the p-values does not exceed the 0.05 threshold, so even in our ultimate model, there is not enough evidence to reject the conditional PTA on a 5% level indicating that estimations perform better than that of the GPA (both UB and LB). Regarding the aggregated coefficient sizes, they are significantly different from zero suggesting that math grades are more affected by family disruption compared to GPAs (-0.09 versus [-0.011; -0.026]).

# Chapter 5. Limitations

In the past, the model was heavily applied for test scores for a reason, as those much more capable of isolating the school fixed effect, given that the tests are similar for all (Keller, 2015). As opposed to that, GPAs can bear large variances due to the difference in expectations, and level of education. In other words, they are normalized within a class to predetermined distribution, making them difficult to compare them within schools (Cunha & Miller, 2014). I used GPAs instead of “competence” math and reading scores simply because half of the students who took the test in 2006, did not have a matching ID in the 2008 test scores. In addition, in the restricted sample of test scores, cases of separations that occurred between 2006 and 2008 lacked in size, meaning that the results were not representative for out-of-sample inferences. All in all, the GPAs and math grades rather than test scores, are likely to be biased even though I included school covariates.

Further, we can imagine other confounding variables such as the quality of the teachers, the size of the classroom, the educational aspiration within the classroom, etc., which were not available in the underlying data to offset the school fixed effect in GPAs. Besides, specifications are describing the ability of students such as whether they have disabilities, or whether they need special educational needs are missing from the longitudinal dataset. By observing an aggregated score such as GPA and math grades, we do not account for the variances in-between the subjects, therefore a supplementary paper should include several additional terms to extend the general model to the joint distribution of multiple subject grades, given teacher, and classroom effects. Due to the lack of extreme cases of misbehavior, such as grade repeat, suspension, or excel, my study does not empirically cover the effect of family disruption on the likelihood of academic failure. Although HOME scores serve as a great proxy of cognitive and emotional abilities, it was only constructed in 2006, thus we assumed that it is a time-invariant (fixed-effect) variable.

The simplified interpretation of parental separation aggregates all scenarios, let it be divorce, separation, moving abroad, decease, prison, etc., which allows for bias in the coefficients. The exact separation of children who were separated from their parents is unknown from the provided data, only the year of separation which may underestimate the coefficients given that the effect is presumably milder as time elapses. Additionally, the data does not allow us to measure the conflict rate or behavior within the household, hence, without estimating the separation probabilities, the variance is incorporated in our results. Lastly, there may arise systemic errors when respondents were asked about sensitive (or subjective) information such as income, ethnical identity, effort, etc.

Even though the staggered adoption solves common issues of the DID, it also has its limitations. To name a few, firstly, our cross-sectional dataset at hand does not allow for more than two pre-treatment periods, thus the t-test on the validity of common trend assumption may have larger confidence intervals as it was seen on the event-study charts. Secondly, the set of used covariates covers most of the in-sample variation, nonetheless, not all explanatory variables mentioned in the literature review were implemented as it was supposed to be due to the lack of relevant cases, such as the recent school changes, and the types of separation (e.g., death, prison, etc.). Thirdly, undetected biases may drive the treatment effect, let it be the change in the law of divorce or an unexpected event that increases the chances of being exposed to treatment because in our estimation we assumed that children do not anticipate the occurrence of treatment. We did so because the short time interval did not allow us to implement such adjustments. Fourthly, we hypothesized that the treatment is irreversible after it had happened, but we can imagine scenarios where parents are separated from each other (divorce or temporary moving due to work) and later rejoin the household. Fifthly, the cohort assembled in the HLCS study is not fully homogeneous, therefore, the conditional PTA only partially holds in all cross sections of my detailed estimations. Sixthly, our model consisted solely linear variables, therefore we assumed that the relationship between the variables does not take polynomial functions.

# Chapter 6. Conclusion

# References

Amato, P. R. (1987a). Children’s Reactions to Parental Separation and Divorce: The Views of Children and Custodial Mothers. *Australian Journal of Social Issues*, *22*(4), 610–623. https://doi.org/10.1002/j.1839-4655.1987.tb00847.x

Amato, P. R. (1987b). Family Processes in One-Parent, Stepparent, and Intact Families: The Child’s Point of View. *Journal of Marriage and the Family*, *49*(2), 327. https://doi.org/10.2307/352303

Amato, P. R., & Anthony, C. J. (2014). Estimating the effects of parental divorce and death with fixed effects models. *Journal of Marriage and Family*, *76*(2), 370–386. https://doi.org/10.1111/jomf.12100

Amato, P. R., & Keith, B. (1991). Parental Divorce and Adult Well-Being: A Meta-Analysis. *Journal of Marriage and the Family*, *53*(1), 43. https://doi.org/10.2307/353132

Astone, N. M., & McLanahan, S. S. (1991). Family Structure, Parental Practices and High School Completion. *American Sociological Review*, *56*(3), 309. https://doi.org/10.2307/2096106

Astone, N. M., & McLanahan, S. S. (1994). Family structure, residential mobility, and school dropout: A research note. *Demography*, *31*(4), 575–584. https://doi.org/10.2307/2061791

Behrman, J. R., & Rosenzweig, M. R. (2004). Returns to birthweight. *Review of Economics and Statistics*, *86*(2), 586–601. https://doi.org/10.1162/003465304323031139

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How Much Should We Trust Differences-In-Differences Estimates? *The Quarterly Journal of Economics*, *119*(1), 249–275. https://doi.org/10.1162/003355304772839588

Biblarz, T. J., Gottainer, G., Journal, S., & May, N. (2016). *Family Structure and Children ’ s Success : A Comparison of Widowed and Divorced Single- Mother Families Published by : National Council on Family Relations Stable URL : http://www.jstor.org/stable/1566757 Family Structure and Children ’ s Success : A Com*. *62*(2), 533–548.

Bierens, B. Y. H. J., & Ploberger, W. (2016). *Asymptotic Theory of Integrated Conditional Moment Tests Author ( s ): Herman J . Bierens and Werner Ploberger Published by : The Econometric Society Stable URL : http://www.jstor.org/stable/2171881 Accessed : 04-06-2016 13 : 19 UTC Your use of the JSTOR* . *65*(5), 1129–1151.

Bradley, R. H., Corwyn, R. F., Caldwell, B. M., Whiteside-Mansell, L., Wasserman, G. A., & Mink, I. T. (2000). Measuring the Home Environments of Children in Early Adolescence. *Journal of Research on Adolescence*, *10*(3), 247–288. https://doi.org/10.1207/SJRA1003\_1

Callaway, B., & Sant’anna, P. H. C. (2019). *Supplementary Appendix: Difference-in-Differences with Multiple Time Periods*. *1*(0). https://economics.uchicago.edu/sites/economics.uchicago.edu/files/uploads/PDF/Callaway\_SantAnna\_2019\_Appendix.pdf

Callaway, B., & Sant’Anna, P. H. C. (2018). *Difference-in-Differences with Multiple Time Periods*. 1–37. http://arxiv.org/abs/1803.09015

Canon, M. E. (2012). Out-of-School Suspensions and Parental Involvement in Children’s Education. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2172022

Cappellari, L. (2004). *High School Types , Academic Performance and Early Labour Market Outcomes*. *1048*.

Case, A. C., Lubotsky, D. H., & Paxson, C. H. (2005). Economic Status and Health in Childhood: The Origins of the Gradient. *SSRN Electronic Journal*, *March*. https://doi.org/10.2139/ssrn.275152

Cross, C. J. (2020). Racial/Ethnic Differences in the Association Between Family Structure and Children’s Education. *Journal of Marriage and Family*, *82*(2), 691–712. https://doi.org/10.1111/jomf.12625

Cunha, J. M., & Miller, T. (2014). Measuring value-added in higher education: Possibilities and limitations in the use of administrative data. *Economics of Education Review*, *42*, 64–77. https://doi.org/10.1016/j.econedurev.2014.06.001

Downey, D. B. (1994). The School Performance of Children From Single-Mother and Single-Father Families:: Economic or Interpersonal Deprivation? *Journal of Family Issues*, *15*(1), 129–147. https://doi.org/10.1177/019251394015001006

Featherman, D. L., & Hauser, R. M. (1978). *Opportunity and change*. Academic Press.

Hargreaves, M. B. (1992). Learning Under Stress: Children of Single Parents and the Schools. In *Scarecrow Press (1992)*. https://philpapers.org/rec/HARLUS-2

Hoxby, C. M. (2004). *Achievement in Charter Schools and Regular Public Schools in the United States: Understanding the Differences.* https://www.semanticscholar.org/paper/Achievement-in-Charter-Schools-and-Regular-Public-Schools-Hoxby/cfcea9b9fed7ee38f271ff28162e98977c0a4dd3

Ingersoll, G. M., Scamman, J. P., & Eckerling, W. D. (1989). Geographic Mobility and Student Achievement in an Urban Setting. *Educational Evaluation and Policy Analysis*, *11*(2), 143–149. https://doi.org/10.3102/01623737011002143

Jeynes, W. H. (1999). The Role of Family Residential Mobility in Explaining the Lower Academic Achievement of High School Children from Reconstituted Families. *Journal of Divorce & Remarriage*, *32*(1–2), 123–143. https://doi.org/10.1300/J087v32n01\_08

Jeynes, W. H. (2000). The effects of several of the most common family structures on the academic achievement of eighth graders. *Marriage and Family Review*, *30*(1–2), 73–97. https://doi.org/10.1300/J002v30n01\_06

Jeynes, W. H. (2005). Effects of Parental Involvement and Family Structure on the Academic Achievement of Adolescents. *Marriage & Family Review*, *37*(3), 99–116. https://doi.org/10.1300/J002v37n03\_06

Judit, M., & Lívia, M. (2015). *Család- és háztartásszerkezet*. 153–170.

Keller, T. (2015). The Power of Family? The Change in Academic Achievement after Breakdown of the Biological Family. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2697532

Kertesi, G., & Kézdi, G. (2011). The Roma/non-Roma test score gap in Hungary. *American Economic Review*, *101*(3), 519–525. https://doi.org/10.1257/aer.101.3.519

Kim, H. J. (2004). Family resources and children’s academic performance. *Children and Youth Services Review*, *26*(6), 529–536. https://doi.org/10.1016/j.childyouth.2004.02.008

Központi Statisztikai Hivatal. (2018). *6. A háztartások és a családok adatai*. https://www.ksh.hu/docs/hun/xftp/idoszaki/mikrocenzus2016/mikrocenzus\_2016\_6.pdf

Központi Statisztikai Hivatal. (2019). A válás demográfiai jellemzői. *Statisztikai Tükör*.

Krein, S. F., & Beller, A. H. (1988). Educational attainment of children from single-parent families: Differences by exposure, gender, and race. *Demography*, *25*(2), 221–234. https://doi.org/10.2307/2061290

Kurtz, L. (1994). Teacher Assessments of Coping Styles in Children of Divorce. *American Educational Research Association*.

Lubienski, S. T., & Lubienski, C. (2006). Mathematics Data. *American Educational Research Journal*, *43*(4), 651–698.

Madden-Derdich, D. A., & Leonard, S. A. (2000). Parental role identity and fathers’ involvement in coparental interaction after divorce: Fathers’ perspectives. *Family Relations*, *49*(3), 311–318. https://doi.org/10.1111/j.1741-3729.2000.00311.x

Manning, W. D., & Lamb, K. A. (2003). Adolescent Well-Being in Cohabiting, Married, and Single-Parent Families. *Journal of Marriage and Family*, *65*(4), 876–893. https://doi.org/10.1111/j.1741-3737.2003.00876.x

Mechanic, D., & Hansell, S. (1989). Divorce, family conflict, and adolescents’ well-being. *Journal of Health and Social Behavior*, *30*(1), 105–116. https://doi.org/10.2307/2136916

Monostori, J. (2019). Egyszülős családok és politikák Magyarországon és Európában. *Demográfia*, *62*(1), 5–41. https://doi.org/10.21543/dem.62.1.1

Monostori, J., & Murinkó, L. (2015). Family and Household Structure. *Encyclopedia of Sociology. Volume 2*, 151–168.

Myers, D. E., Milne, A. M., Baker, K., & Ginsburg, A. (1987). Student Discipline and High School Performance. *Sociology of Education*, *60*(1), 18. https://doi.org/10.2307/2112616

Newhouse, D. L., & Beegle, K. (2011). The Effect of School Type on Academic Achievement: Evidence from Indonesia. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.753544

Nilsen, S. A., Breivik, K., Wold, B., Askeland, K. G., Sivertsen, B., Hysing, M., & Bøe, T. (2020). Divorce and adolescent academic achievement: Heterogeneity in the associations by parental education. *PLoS ONE*, *15*(3), 1–22. https://doi.org/10.1371/journal.pone.0229183

*NLSY79 Child and Young Adult Data User’s Guide 2002 Appendix A.* (2004). Centre for Human Resource Research. The Ohio State University. https://www.nlsinfo.org/pub/usersvc/Child-Young-Adult/2006ChildYA-DataUsersGuide.pdf

OECD. (2016). Family Size and Composition. *OECD Family Database*, 6–12. http://www.oecd.org/els/family/SF\_1\_1\_Family\_size\_and\_composition.pdf

OECD. (2017). *SF3 . 2 . Family dissolution and children*. *December*, 1–5. http://www.oecd.org/els/family/SF\_3\_2\_Family\_dissolution\_children.pdf

Pong, S.-L., & Ju, D.-B. (2000). The Effects of Change in Family Structure and Income on Dropping Out of Middle and High School. *Journal of Family Issues*, *21*(2), 147–169. https://doi.org/10.1177/019251300021002001

Shinoda, K. S. (2011). *Children of Divorce: The Impact on Classroom Behavior*.

South, S. J., Crowder, K. D., & Trent, K. (1998). Children’s residential mobility and neighborhood environment following parental divorce and remarriage. *Social Forces*, *77*(2), 667–693. https://doi.org/10.1093/sf/77.2.667

Teachman, J. D. (1987). Family Background, Educational Resources, and Educational Attainment. *American Sociological Review*, *52*(4), 548. https://doi.org/10.2307/2095300

Teyber, E. (2001). *Helping Children Cope With Divorce*. John Wiley & Sons, Inc. (US).

Thomson, E., Hanson, T. L., & Mc Lanahan, S. S. (1994). Family structure and child well-being: Economic resources vs.parental behaviors. *Social Forces*, *73*(1), 221–242. https://doi.org/10.1093/sf/73.1.221

Tucker, C. J., & Long, L. (2018). *" Moving On ": Residential Mobility and Children ’ s School Lives Author ( s ): C . Jack Tucker , Jonathan Marx and Larry Long Published by : American Sociological Association Stable URL : https://www.jstor.org/stable/2673244 American Sociological Associa*. *71*(2), 111–129.

Wallerstein, J. S., & Corbin, S. B. (1989). Daughters of divorce: Report from a ten-year follow-up. *American Journal of Orthopsychiatry*, *59*(4), 593–604. https://doi.org/10.1111/j.1939-0025.1989.tb02749.x

Wallerstein, J. S., & Kelly, J. (1980). Surviving the Breakup: How Children and Parents Cope with Divorce. In *New York: Basic Books*.

Wing, C., Simon, K., & Bello-gomez, R. A. (2018). *Designing Difference in Difference Studies : Best Practices for Public Health Policy Research*.

Wright, S. P., Horn, S. P., & Sanders, W. L. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education*, *11*(1), 57–67. https://doi.org/10.1023/a:1007999204543

Yongmin Sun, & Yuanzhang Li. (2009). Postdivorce Family Stability and Changes in Adolescents’ Academic Performance. *Journal of Family Issues*, *30*(11), 1527–1555. https://doi.org/10.1177/0192513X09339022

# Appendix

|  |  |  |  |
| --- | --- | --- | --- |
|  | Family structure | | |
| Gender | Intact | Non-intact | Total |
| Male | 54.23% | 53.52% | 53.97% |
| Female | 45.77% | 46.48% | 46.03% |
| Total | 100.00% | 100.00% | 100.00% |

**Appendix 3.1.** Distribution of gender across family types (intact = two-parent biological family, s = any other disrupted family types).

**Appendix 3.2.** Distribution of children across family types to whom at least one parent has a Roma-origin.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Family structure | | |
| Roma-origin | Intact | Non-intact | Total |
| 0 | 90.61% | 94.11% | 91.86% |
| 1 | 8.77% | 5.45% | 7.57% |
| <NA> | 0.61% | 0.44% | 0.57% |
| Total | 100.00% | 100.00% | 100.00% |

**Appendix 3.3.** Distribution of family structures

|  |  |  |
| --- | --- | --- |
|  | Frequency | % Valid |
| Intact | 6530 | 65.78 |
| Non-intact | 3397 | 34.22 |
| <NA> | 95 | NA |
| Total | 10022 | 100.00 |

|  |  |  |
| --- | --- | --- |
|  | Frequency | % Valid |
| Alone (no parents) | 57 | 0.58 |
| Cohabiting stepfather | 385 | 3.92 |
| Remarried stepfather | 355 | 3.61 |
| Single-father foster | 14 | 0.14 |
| Single father | 230 | 2.34 |
| Single mother | 1945 | 19.78 |
| Single-mother foster | 78 | 0.79 |
| Stepmother | 131 | 1.33 |
| Two-parent foster | 106 | 1.08 |
| Two-parent biological | 6530 | 66.42 |
| <NA> | 191 | NA |
| Total | 10022 | 100.00 |

**Appendix 3.3.** Further breakdown of structures

**Appendix 3.4.** Distribution of parental separation types

|  |  |  |
| --- | --- | --- |
|  | Frequency | % Valid |
| Death | 60 | 2.44 |
| Divorce | 2249 | 91.39 |
| Other | 152 | 6.18 |
| <NA> (two-parent) | 7561 | NA |
| Total | 10022 | 100 |

**Appendix 3.5.** Age of children at the parental separation

|  |  |
| --- | --- |
| Age at maternal separation | Age at paternal separation |
| Min.: 0.000 | Min.: 0.000 |
| 1st Qu.: 3.000 | 1st Qu.: 3.000 |
| Median: 8.000 | Median: 7.000 |
| Mean: 7.641 | Mean: 6.799 |
| 3rd Qu.: 12.000 | 3rd Qu.: 11.000 |
| Max.: 18.000 | Max.: 20.000 |

**Appendix 3.6.** School grades and competence test scores across family types

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Mean GPAs (on a scale to 5) | | | | | | Competence test scores (max. 500) | |
|  | GPA | Math | Grammar | Literature | Behaviour | Diligence | Math | Reading |
| Intact | 3.79 | 3.32 | 3.53 | 3.74 | 4.18 | 3.85 | 480.93 | 465.77 |
| Non-intact | 3.61 | 3.08 | 3.37 | 3.56 | 3.98 | 3.66 | 465.24 | 452.45 |

**Appendix 3.7.** Parental income and the propensity to consume (monthly net amounts) in HUF.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Propensity to consume | Maternal Salary | Paternal Salary |
| Intact | 160,090.4 | 84,560.87 | 104,398.78 |
| Non-intact | 128,988.43 | 88,394.28 | 101,421.43 |

**Appendix 3.8.** Scales to measure the emotional support and cognitive stimulation of children. Home Score equals to the aggregated cognitive and emotional scores.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Home Score | Cognitive Score | Emotional Score |
| Intact | 183.8 | 82.08 | 101.91 |
| Non-intact | 166.65 | 75.39 | 91.52 |

**Appendix 3.9.** Grade repeats across family types

Grade repeats before 4th grade

|  |  |  |  |
| --- | --- | --- | --- |
|  | Family structure | | |
| No. of grade repeats | Intact | Non-intact | Total |
| 0 | 93.95% | 90.99% | 92.94% |
| 1 | 5.38% | 7.86% | 6.23% |
| 2 | 0.47% | 0.85% | 0.60% |
| 3 or more | 0.06% | 0.06% | 0.06% |
| <NA> | 0.14% | 0.24% | 0.18% |
| Total | 100.00% | 100.00% | 100.00% |

**Appendix 3.9.** Grade repeats across family types

Grade repeats between 5th and 8th grade

|  |  |  |  |
| --- | --- | --- | --- |
|  | Family structure | | |
| No. of grade repeats | Intact | Non-intact | Total |
| 0 | 95.38% | 92.55% | 94.39% |
| 1 | 3.74% | 6.30% | 4.62% |
| 2 | 0.44% | 0.53% | 0.48% |
| 3 or more | 0.08% | 0.21% | 0.12% |
| <NA> | 0.37% | 0.41% | 0.39% |
| Total | 100.00% | 100.00% | 100.00% |

**Appendix 3.10.** Expels across family types

|  |  |  |  |
| --- | --- | --- | --- |
|  | Family structure | | |
| No. of expels | Intact | Non-intact | Total |
| 0 | 99.79% | 99.32% | 99.63% |
| 1 | 0.21% | 0.65% | 0.36% |
| 3 | 0.00% | 0.03% | 0.01% |
| Total | 100.00% | 100.00% | 100.00% |

**Appendix 3.11.** Dropouts (leaving due to low grades) across family types

|  |  |  |  |
| --- | --- | --- | --- |
|  | Family structure | | |
| No. of dropouts | Intact | Non-intact | Total |
| 0 | 98.79% | 97.38% | 98.31% |
| 1 | 1.13% | 2.38% | 1.56% |
| 2 | 0.06% | 0.18% | 0.10% |
| 3 | 0.02% | 0.06% | 0.03% |
| Total | 100.00% | 100.00% | 100.00% |

**Appendix 3.12.** Measures of parental school engagement (Overall score = Sum of first 3 columns; the higher the better)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Attend parent meetings | Talk to teachers | Study w/ the child | Overall Score |
| Intact | -1.24 | -1.74 | -2.25 | -5.23 |
| Non-intact | -1.35 | -1.78 | -2.4 | -5.54 |

**Appendix 3.13.** Monthly average parental school-related spending across family types in HUF.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Textbooks | Transportation | Extra classes | School trips | Total |
| Intact | 26,790.42 | 2,111.00 | 1,109.24 | 3,763.29 | 33,773.95 |
| Non-intact | 19,559.32 | 1,769.54 | 780.85 | 3,564.46 | 25,674.17 |

**Appendix 3.14.** Number of school changes due to moving.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Family structure | | |
| No. of school change due to moving | Intact | Non-intact | Total |
| 0 | 93.29% | 83.34% | 89.93% |
| 1 | 5.68% | 12.04% | 7.81% |
| 2 | 0.84% | 3.41% | 1.72% |
| 3 | 0.17% | 0.91% | 0.43% |
| 4 | 0.02% | 0.29% | 0.11% |
| Total | 100% | 100% | 100% |

**Description of variables**

**GPA (float):** The arithmetic average of all subject grades ranging from 1 (fail) to 5 (excellent).

**Roma-origin (binary):** The caretaker mother and father of the respondents were asked whether they have a Roma-origin, and if at least one of them identified themselves as a Roma (primarily or secondarily), the dummy of one was assigned to the categorical variable.

**Health status (categorical):** Respondents were asked the following question with the options listed below:

*How would you rate the state of your health?*

1 – Excellent

2 – Good

3 – satisfactory or

4 – bad?

**Birthweight (binary):** Reported birthweight in grams. As in the cited literatures, I also used the threshold of 2,500 grams as a measure of low birthweight. Therefore, the variable is a dummy construction which takes one if the respondent born with a low birthweight (below 2,500 grams).

**No. Minors (integer):** Respondent were asked on the number of minors in different age groups (0-3, 4-6, 7-14, 15-18 of age) within the household. The variable aggregates all age groups to determine the minor cohabitants.

**School Type (categorical):** Respondents were asked the following question with the options listed below:

*What type of school does he/she go to?*

1 – general school

2 – vocational (formerly: skilled worker) training school

3 – vocational secondary school

4 – 4 year academic secondary school

5 – 6 year academic secondary school

6 – 8 year academic secondary school

7 – other school

*Inputs were reduced to three distinct categories: 1 - vocational or equivalent 2 - high school 3 - other. There were no students who attended general school.*

**School Maintainer (categorical):** Respondents were asked the following question with the options listed below:

*What is the school’s educational provider?*

1 – the state or a local authority,

2 – the church,

3 – it is a private school run by a foundation or other source

**Household Income (float):** Aggregated reported net monthly income (HUF) of the caretakers in the household on a logarithmic scale.

**Welfare (categorical):** Respondents were asked the following question with the options listed below:

*Compared to other families, how well does your family live?*

1 – They get on very badly, have to give up many things.

2 – They can only get on badly.

3 – They live on an average level.

4 – They live better than the average.

5 – They live very well, can allow almost everything.

**Maternal Degree (categorical):** Female caretakers were asked the report their educational attainment on the basis of the following categories:

00 – general school age or pre-school age

01 – less than lower secondary school

02 – lower secondary school

03 – basic vocational education (uncertified)

04 – vocational certificate

05 –maturity certificate in vocational secondary school

06 – maturity certificate in academic secondary school

07 – post secondary, non-tertiary education

08 – college level or BA

09 – university level or MA

10 – doctoral or equivalent level (PhD, DLA)

**Parental School Involvement (PSI) (integer):** We can distinguish two variables (1) Parental school engagement rate (frequency of parent-teacher talk and attending parent conferences), and the (2) monthly aggregated school investments (HUF) on textbooks or equipment, transportation costs (e.g., public transportation passes, petrol to personal vehicle, etc.), school trips and extra classes on a logarithmic scale. The former indicator is the sum of the given categories given in the two questions below multiplied by negative one, so the higher, the better the engagement.

*How often did they attend ..................’s parent-teacher conferences since 5th grade?*

1 – almost always,

2 – usually,

3 – a few times, or

4 – almost never attend them

*Apart from parent-teacher conferences, how often have they seen the teachers since ................... went to 5th grade?*

1 – Often,

2 – rarely, or

3 – never?

**Residential Mobility (integer):** Respondents were asked to list all the schools he/she attended until 2006 and why did he/she switch from one to another. The number of school changes due to moving were summed up to determine the mobility rate of the student in the pre-treatment period.

1. Children who experienced a divorce before 14 and possibly additional family transition but the structure was stable between the waves. [↑](#footnote-ref-1)
2. Children who experienced a divorce before 14 and the transition was in-between the waves. [↑](#footnote-ref-2)