THE ROCKWOOL FOUNDATION
Research

THE EFFECTS OF PARENTAL UNION DISSOLUTION ON CHILDREN'S TEST SCORES

MATHILDE LUND HOLM PETER FALLESEN ESKIL HEINESEN

STUDY PAPER 185 JANUARY 2023

The effects of parental union dissolution on children's test scores

Study Paper No. 185

Published by:

Address:
The ROCKWOOL Foundation Research Unit
Ny Kongensgade 6
1472 Copenhagen, Denmark

Telephone +45 33 34 48 00
E-mail: kontakt@rff.dk
https://www.rockwoolfonden.dk/en

January 2023

THE EFFECTS OF PARENTAL UNION DISSOLUTION ON CHILDREN'S TEST SCORES*

Mathilde Lund Holm[1,2] (mlh@econ.ku.dk)
Peter Fallesen[2,3] (peter.fallesen@sofi.su.se)
Eskil Heinesen[2] (esh@rff.dk)

[1] *Department of Economics, University of Copenhagen*
[2] *ROCKWOOL Foundation*
[3] *Swedish Institute for Social Research, Stockholm University*

ABSTRACT

Many children experience parental union dissolution during childhood. This study provides evidence on the immediate and long-term effects of parental separation and union dissolution on children's test scores. We use population data on parents moving out of the joint home and national school-administered low-stakes test. Using a staggered event-study design, we find dynamic long-term negative effects on test scores. The decline in test scores originates from the middle of the skill distribution. Further, we demonstrate plausible indications of an immediate negative effect of parental separation on children's test scores using a regression discontinuity design.

Keywords: academic achievement, children, distributional effects, divorce, test scores

JEL: J12, J13, J24

---

* Corresponding author: Peter Fallesen. The study received funding from the ROCKWOOL Foundation (Grant no. 1207). Peter Fallesen also acknowledges funding from Riksbankens Jubileumsfond (Grant no. P20-0270). We thank Mette Ejrnæs, Maria Stanfors, Mar Espadafor, Alicia Sierra, and Ellen Greaves for excellent suggestions and comments. We are furthermore grateful for comments received at presentations at the Population Association of America annual meeting, the Copenhagen Education Network, the ROCKWOOL Foundation, University of Copenhagen, and the European University Institute.

1. INTRODUCTION

In most OECD countries the divorce rate has increased substantially since 1970 reaching a plateau around the turn of the century (Wagner 2020). In the US, around 40 to 50 percent of all marriages end in divorce (Kennedy and Ruggles 2014). The divorce rate in Denmark is comparable to the US and varied from 2000 to 2018 between 43 and 53 percent.[1] Both Denmark and the US had higher divorce rates in 2020 compared to the OECD average (OECD 2022). In Denmark, more than 30 percent of children can expect to experience a parental divorce or union dissolution during their childhood. Such shocks to family life may drastically alter the home environment and impose new constraints on parents' time and economic resources.

As divorce has become more common in society, a large literature examining different aspects of divorce has grown. One focus of the divorce literature has been on how divorce laws affect the frequency of divorce (e.g., Fallesen 2021; Kabátek 2019; Kneip, Bauer, and Reinhold 2014; Lee 2013; Wolfers 2006). Another area of interest has been on how divorce and parental separation affect the children of the household in terms of both short– and long-run outcomes, such as human capital formation, behavioral problems, mental health, or income levels in adulthood. Children of divorced parents reported poorer mental wellbeing and more behavioral problems (e.g., Fallesen and Gähler 2020a; Strohschein 2005), had lower academic achievement (e.g., Kim 2011; Sigle-Rushton et al. 2014), and lower academic attainment (Heinesen 2019; Laird, Nielsen, and Nielsen 2020). Yet, due to the complex and non-random nature of divorce (Lyngstad and Jalovaara 2010), strong causal designs have been elusive with most studies adjusting for confounding through either conditioning on observables, matching strategies, latent variable modelling, and sibling fixed effect

[1] 20-year marriage survival rate reported by Statistics Denmark.

designs (see McLanahan, Tach, and Schneider 2013 for review). Thus, selection issues may persist, especially in the light of evidence on pre-divorce adjustments (e.g., Sun and Li 2001), selection not only into divorce but also into timing of divorce (e.g., Brüderl and Kalter 2001; Fallesen and Breen 2016), and the role of time-and-country specific divorce norms and opportunities (de Graaf and Kalmijn 2006; Gruber 2004; Kreidl, Štípková, and Hubatková 2017; Piketty 2003).

In this paper we use a novel data source (the results from nationwide national reading comprehension tests of all Danish children in public schools taken at grades 2, 4, 6, and 8) as well as information on parental separation captured at the date the parents moved apart to study the effect parental separation and union dissolution/divorce on children's cognitive development.[2] The national tests are low stakes and have no consequences for the child's future. For this reason, parents might be less likely to try to affect the result through investments such as tutoring—thus, the national tests might be an accurate measure of child ability. Further, given the test's low stakes, parents are also not likely to time a separation relative to the test (i.e., no incentive for bunching). We provide important evidence on three aspects of the effect of parental union dissolution on children's educational achievement. First, using a within-person event-study design with stacked regressions, we can account for selection into experiencing parental divorce and estimate effects that may depend on time of treatment and time since treatment. Applying the recent dynamic difference-in-differences (DID) approach by Callaway and Sant'Anna (2021) allows us to bolster a causal and dynamic interpretation of the impact of parental separation and union dissolution on children's educational performance. Second, we examine heterogeneity across the distribution of

[2] In a Danish context it is not important to distinguish between married and non-married parents who live together. Often parents are not married when the first child is born, but then become married some years later. Björklund and Sundström (2006) and Sigle-Rushton et al. (2014) analyzing data for Sweden and Norway, respectively, also focus on separation irrespective of prior marital status.

test score performance by estimating counterfactual distributions (Chernozhukov, Fernández-Val, and Melly 2013), demonstrating that the effect of parental union dissolution is heterogenous across the test score distribution. Last, we can recover the effect of the discrete event of parents moving apart on test scores through a regression discontinuity design (RDD). To our knowledge, this paper is the first to analyze effects of parental separation on test scores using dynamic DID models taking account of individual fixed effects, to investigate heterogeneity in these effects across the test score distribution, and to use an RDD to analyze the immediate effect of parents moving apart.

In our event-study analysis we find negative effects of parental union dissolution on children's test scores in the range of 3%-7% of a standard deviation. There are indications of dynamic effects that increase by time since union dissolution. The results are robust to applying recent advances in the estimation of DID models (Callaway and Sant'Anna 2021). The effect of union dissolution on children's test score is driven by children in the middle 50 % of the test score distribution. Further, from our RDD analysis, we find at the separation discontinuity a decrease in test scores of 3%-4% of a standard deviation. There is no indication of parents strategically changing separation timing around the test time. Our findings are in accordance with findings from the medical literature of increased stress response and delayed maturation in children following parental separation (Gerra et al. 1993; Sheppard, Garcia, and Sear 2015; Vezzetti 2016). In total, we provide evidence that parental separation and union dissolution has statistically and substantially significant negative effects on children's educational performance in both the short and the long run.

2. SEPARATION, UNION DISSOLUTION, AND CHILDREN'S DEVELOPMENT

Why Separation and Union Dissolution Matters for Children

A large literature have documented how children who experience parental union dissolution are disadvantaged across a host of outcome measures, but have also cautioned that it is not random

which families divorce (see Amato 1993; 2000; 2010; Lyngstad and Jalovaara 2010 for reviews). In this section we will first lay out the likely causal pathways through which parental separation and union dissolution may affect children's educational performance, then discuss how the literature has examined this question, and lastly discuss how to define the counterfactual state for a child experiencing separation and union dissolution.

Underlying mechanisms

There is a long tradition in the family and public health literature to consider guardian separation and divorce as an adverse childhood experiences (ACE), in children's lives (e.g., Brown et al. 2009; Crouch et al. 2019; Lew and Xian 2019).[3] In the medical literature, children who experience parental separation and union dissolution are found to slow down in their biological maturation following the event (Gerra et al. 1993; Sheppard, Garcia, and Sear 2015) and report increased levels of stress (see Vezzetti 2016 for a recent review). Both responses are likely to inhibit learning and/or performance in the educational system. If learning is affected, the consequences of psychosocial trauma of a parental separation may fester assuming a dynamic model of learning (cf. Cunha and Heckman 2007).

Second, from a pure economies of scale perspective, divorce will (most often) lead to reduced economic resources available for children.[4] Exposure to a low income childhood environment is associated with lower educational attainment (Lesner 2018), perhaps due to fewer resources for parents to invest in their children (cf. Cunha and Heckman 2007). Another resource that becomes

---

[3] Although parental divorce was not considered in Felitti et al. (1998) seminal study of ACEs.

[4] The form of this mechanism depends on the extent to which especially women increase labor supply following the end of a relationship (Bargain et al. 2012; Bonnet, Garbinti, and Solaz 2021). Other studies have found that women pay a higher and continuous price for divorcing (Smock, Manning, and Gupta 1999; Leopold 2018; Raz-Yurovich 2013), and that welfare state arrangements may play a moderating role (Uunk 2004).

more scarce following parental divorce is parents' time with children. Fallesen and Gähler (2020b) show that the time parents spend on developmental tasks with their children declines following a union dissolution, whereas time spent on non-developmental (practical) tasks remains the same. Thus, a union dissolution likely leads parents to have fewer economic resources and time to invest in children. Insofar that the dividends of such investments are compounded (such that skill begets skill), the negative consequence of parental union dissolution may increase over time (i.e., the effect is dynamic).

Parental union dissolution may not affect all children equally. Children's ability to navigate and lessen negative impacts of adverse events are often referred to as resilience (see Zolkoski and Bullock 2012 for review). A key predictor of resilience is intelligence (Cederblad et al. 1995; Katz and Gottman 1997; Friborg et al. 2005; Condly 2006; Alvord and Grados 2005), which strongly correlates with academic ability, achievement, and test scores. If resilience also buffers negative effects of parental separation and union dissolution, it will indicate heterogenous effects across children's educational ability. Further, the level of parental conflict prior to separation may also be related to heterogeneity in separation effects; a high level of conflict may reduce the negative effect of separation (Clark et al. 2015; Garriga and Pennoni 2022) and in turn also already affect children's test scores negatively. Thus, to the extent that low performance is related to poor home environments, a change to the home environment through parental union dissolution may be of little consequence.

Previous work on divorce and educational outcomes in children

The main challenge when studying the impact of divorce, union dissolution, and parental separation on child outcomes is selection and timing effects. Families in which the parents choose to divorce (or separate) are likely to be different from families where the parents stay continuously married

(e.g., Heinesen 2019). Regardless of divorce, these differences are likely to affect child outcomes and consequently, parents who do not divorce or separate may be an inappropriate control group for the parents who do divorce or separate. To overcome the issue of selection, some studies have utilized longitudinal data and aimed to control for relevant confounders through various statistical techniques (e.g., Kim 2011), or have used the timing of the divorce to obtain causal estimates of how divorce and separation affect child outcomes. In this latter setting, older siblings experiencing parental separation as adults (e.g., after age 18) or after a given outcome is measured can serve as a control group for younger siblings experiencing parental separation earlier in life or prior to the outcome is measured. This has commonly been called a sibling fixed effect approach (e.g., Björklund, Ginther, and Sundström 2007; Björklund and Sundström 2006; Heinesen 2019; Laird, Nielsen, and Nielsen 2020; Sigle-Rushton et al. 2014).

The sibling fixed effect methodology controls for constant family characteristics and some family characteristics are, undoubtedly, constant. It seems reasonable, however, to assume that the conflict level or other social problems in the family increases as the time of divorce approaches (see Härkönen, Bernardi, and Boertien 2017 for recent review). Consequently, a set of siblings raised by the same parents might experience quite a different family environment during childhood and this could potentially cause a bias in the estimated effect of separation. Union dissolution is recorded as a discrete event, but likely reflects the consequences of long run deterioration of the parental relationship. A second issue for the sibling fixed effect methodology is if certain child characteristics are the cause of parental separation. For instance, if a younger sibling is born with a disability or is particularly challenged in school this could be a contributing factor to parental separation and affect educational achievement of the child (Kvist, Nielsen, and Simonsen 2013; Mallinson and Elwert 2022). Thus, the identification strategies often employed in the literature (Sigle-Rushton et al. 2014; e.g., Björklund, Ginther, and Sundström 2007; Björklund and

Sundström 2006; Kim 2011) rest on strong and often untestable assumptions. A final issue is potential dynamic effects. If the impact of parental union dissolution increases with time (e.g., leads to lower parental investment in children, which in turn leads to lower returns that become compounded), this would lead to dynamic effects. If this occurs, then the relative differences between the occurrence of separation and measuring the outcome of interest becomes of pivotal importance. Yet, the counterfactual sibling outcome likely will be ill-defined unless the functional form of the dynamic effect is correctly specified in terms of the distance between the treated siblings' outcome and the timing of the union dissolution.

A second strand of literature has focused on the impact of changes in divorce laws, either as a direct shock to parental relationships (e.g., Piketty 2003) or as an exposure measure, comparing people who spent different amounts of their childhood under a more liberal divorce policy regime (Cáceres-Delpiano and Giolito 2012; González and Viitanen 2009; Gruber 2004; Johnson and Mazingo 2000). Whereas divorce law changes likely provide exogenous variation, the designs based on them do not capture the direct (nor the total) causal effects of divorce, because the reforms affect child outcomes through other channels: families can substitute official divorce for continuous separation, the reduction in barriers to and costs of exiting marriage may increase marriage rates, the relative bargaining positions of both genders change, and remarriage rates and family complexity may increase (see Gruber 2004; Piketty 2003 for further discussion). In addition, an increasing number of children are born to parents who never marry but instead cohabit, especially in the Nordic countries generally known for liberal family policies that equate cohabiting and married parental rights closely (Perelli-Harris and Gassen 2012). For Denmark alone, 54 percent of children was born out of wedlock in 2020. In that case, changes to divorce law (and divorce as a phenomenon) affects only a subset of the underlying social phenomenon of actual interest—the dissolution of the parental co-residing relationship.

Further, whereas studying the impact of divorce reforms that occurred in 1960s-1970s provides invaluable evidence on the historical consequences of those reforms, they also represent studies of divorces occurring at a very different historical margin than the margin likely to be relevant today. As Piketty (2003) noted, the French couples he studied who became able to divorce following the 1975 introduction of a no-fault, mutual-consent divorce law, likely were couples who had very high-conflict relationships. This in turn may explain why he finds little consequence of the reform on children's outcomes. Presently, divorce has become substantially more common, and the normative barrier for divorcing has likely been lowered. In descriptive work, de Graaf and Kalmijn (2006) document that for the Netherlands, divorce motives became less severe as divorce became more common. That is, the married couples that divorced following the reforms in the 1960s-1970s likely had on average more problematic home environments than couples who divorced or dissolved their union in later periods after divorce became more common and normalized. Thus, children growing up in problematic home environments may not react to (or may even benefit from) a parental union dissolution, whereas children growing up in less problematic home environment where the parents still dissolve their union may experience substantial decline in the quality of the home environment (see also Clark et al. 2015).

Defining the Counterfactual

Studying the effect of parental separation and union dissolution raises a more fundamental question: what is the counterfactual treatment? If we consider the case where one partner meets a new person randomly, through for example changes to their pool of colleagues in the workplace (McKinnish 2004; 2007; Svarer 2007), this counterfactual is well-defined conditional on people's valuation of outside options relative to their present relationship: the counterfactual treatment is the relationship's trajectory had the meeting with a new person never happened. Similarly, any singular unexpected event that causes either partner to leave works in the same way.

Yet, assuming people form unions because they believe they have higher expected utility of being in that union rather than being single, union dissolutions are likely also the result of deteriorating relationship quality lowering present and expected future utility (Becker, Landes, and Michael 1977; Weiss and Willis 1997) and of processes of learning about the actual match quality of the relationship (Fallesen and Breen 2016). In either case, the likely counterfactual treatment to separation at a given time is separation a few years later while the utility of being relationship in the counterfactual state continues to decline.

3. DIVORCE LEGISLATION AND UNION DISSOLUTION IN DENMARK

Divorce legislation in Denmark in the period 2009-2018

Divorce legislation in Denmark was changed by a reform in 2013 (Fallesen 2021; Rosenbeck 2017). From 2009 to 2013 all couples filing for divorce were required to complete a period of legal separation of six months unless aggravating circumstances as infidelity, domestic violence, etc. were present. From July 1 2013, partners were allowed to skip the period of legal separation, given both partners agreed to the divorce. The reform change was found to increase the divorce rate by 10% (Fallesen 2021), an effect similar to related findings from the Netherlands (Kabátek 2019) and South-Korea (Lee 2013). Hence, from July 2013 couples in Denmark could apply either for direct divorce in uncontested cases or still go through a six-months legal separation. For a legal separation period to be valid the couple must be non-cohabiting. To allow time for finding a new residence the couple has 3 months to comply with the demand of separate addresses. It is furthermore compulsory by law to officially register a new address within 5 days after a move. In our study we use the date of move as the time of separation (union dissolution), and this is likely to be prior to the date of formal divorce. Thus, time of separation represents the end of the married union rather than time of formal divorce. By focusing on time of separation (parents moving apart) rather than time of formal

divorce among married couples, we circumvent the shock to divorce timing caused by the 2013 reform, we obtain a definition of date of separation which is consistent across married and unmarried couples, and it is likely to be the important date in terms of effects on children. Nevertheless, the reform might have had effects on divorce and separation decisions of married couples. We take account of this by adjusting for year of test fixed effects in our empirical models.

Cohabiting Unions and Union Dissolution

In recent decades, including those covering the cohorts considered in this study (born 1995-2010), a large share of children is born into and grow up in unmarried unions. 46% of Danish children born 1995-2010 was born out of wedlock, and 14 percent of those still living with both parents at age 8 had parents who remained unmarried.[5] Across the Nordic countries, cohabiting unions in lieu of married unions have long been on the rise (Perelli-Harris et al. 2012; Heuveline and Timberlake 2004), and they have increasingly been bestowed the same legal rights as married unions (Perelli-Harris and Gassen 2012)—especially when those unions involves joint children. However, although parental rights for biological parents (and designated biological parents in the case of adoption or children conceived through medically assisted reproductive techniques) are the same across union type, the full formal framework for ending a union that exist for ending a marriage (i.e., divorce) is not present for unmarried unions. While unmarried unions thus will lack a divorce date, they still have a separation date that can be aligned with the separation date for divorcing married couples, because a separation also occurs either at the same time as a divorce or heralds it.

[5] According to Statistics Denmark: http://statistikbanken.dk/FAM111N

4. RESEARCH DESIGN

It is not random who separates and dissolves their union (selection into union dissolution), nor when it happens (selection into timing). Couples with children who separate have on average shorter education and lower income than those who do not, and also compared to those who separate when their children are older (see Section 5). Further, parents who separate sooner after their child was born are likely less well-matched and had on average been together for a shorter while before becoming parents (Fallesen and Breen 2016). To identify any causal effect of parental separation and union dissolution on children, these two key sources of selection need to be addressed.

Estimating the Long-Term Effect of Dissolution: Event Study DID Models

The dynamic DID models we consider are two-way fixed effects model with fixed effects for individuals (or groups) and time (see, e.g., de Chaisemartin and D'Haultfœuille 2020). The treatment (parental dissolution) is an absorbing state since it may affect children in all subsequent periods. Our models are examples of a staggered adoption design where some children experience parental dissolution early in the school career and others in later grades. The identifying common trend assumption requires that the test score trend for the treatment group in the counterfactual state of non-treatment would have been the same as the trend for the control group. We compare students treated in earlier grades to control groups of not-yet treated (in some analyses combined with a group of never treated). We avoid using earlier treatment groups as control groups for later treated because of the issues related to heterogenous treatment effects discussed in, e.g., de Chaisemartin and D'Haultfœuille (2020) and Goodman-Bacon (2021). Including never treated children in the control group adds statistical power by increasing sample size, but their test score trend may differ from the counterfactual trend of the treatment group because of differences in terms of both

observables and unobservables. Therefore, we present results both with smaller control groups of later treated and larger control groups that also include never treated.

Our outcome variable is test scores in reading comprehension, which are observed in grades 2, 4, 6 and 8.[6] By allowing for at least one pre-treatment observation, we can estimate effects of separation on test scores in grades 4, 6 and 8. To simplify, we define separation as occurring in either grade 2, 4, 6, 8 or 10 according to a "nearest test" criterion. For instance, if a child experiences parental separation in grade 3, this is considered a grade 2 (4) event if the date of separation is closer to the date of the grade 2 (4) test than to the date of the grade 4 (2) test. Separation in "grade 10" means separation within our data period, but at least one year after the test in grade 8.[7] If separation occurs more than one year before the grade 2 test, it is considered to happen "before grade 2". Those who do not experience parental separation within our data period are called "never treated".

Given our nearest test specification, it is reasonable to consider, for instance, grade 2 test scores as a pre-treatment outcome for those experiencing parental separation in grade 4, that is, in the period between about a year prior to the grade 4 test and a year after this test. Thus, the grade 2 test score is measured at least about a year prior to the separation in grade 4. Therefore, it is unlikely that the grade 2 test score is affected by a possible escalation of social problems or conflict in the family leading up to separation (see the discussion in Section 2).[8]

Our nearest test definition of the grade of separation may seem at odds with our RDD analysis discussed in the introduction and in the next subsection, where we estimate a discontinuous drop in test scores at the time of separation. Thus, our nearest test definition ignores such a discontinuity by

---

[6] Test score data are also available for other subjects, but we do not use these data because of a weaker longitudinal structure (math scores in grades 3, 6 and 8; English in grades 4 and 8; physics/chemistry in grade 8 only).
[7] The data which we use to identify separation are updated to more recent years than the test score data.
[8] We report robustness checks that support this presumption.

assuming a common immediate effect in the grade of separation (at event time 0) irrespective of whether the separation occurred before or after the test date. However, in the DID analysis, where the focus is on more long-term effects of separation, it is an advantage to simplify. In the Appendix, we show results for a modified DID model where we allow the effect in the grade of separation to differ by whether separation occurred before or after the test. The results (discussed in more detail in Section 6) indicate a negative effect on test scores at both treatment times, and (in accordance with the RDD analysis) larger effects when separation occurred before the test.

A simple alternative to our nearest test definition, which might seem more in accordance with the RDD results, could be a "before test" definition where, for instance, the grade 4 treatment group consists of those experiencing separation before the grade 4 test but after the grade 2 test. However, here it may be problematic to consider the grade 2 test scores as a pre-treatment outcome for the grade 4 treatment group, since some students in this group experience separation just after the grade 2 test and therefore their grade 2 test scores could be affected by escalation of social problems and conflict in the family leading up to separation (see the discussion in Section 2). With a similar argument, the grade 8 treatment group would not be a valid control group when estimating the effect of separation in grade 6. Thus, with this alternative definition of the grade of treatment, our effective sample would be much smaller, and we could only estimate effects in grades 6 and 8.

Within each model of the form of Eq. (1), the three event time parameters γ_e are estimated for the same sample. However, in the estimation of some of the γ_e parameters, we may include more not-yet-treated groups in the control group. For instance, when $g = 4$ the effect at event time 0 may be estimated including the groups $g = 6$ and $g = 8$ in the control group in addition to the group $g = 10$, or we may use only the group $g = 6$ as control group because it will presumably be the group of not-yet-treated resembling the treatment group $g = 4$ the most making the common trend

assumption more likely to hold. In our empirical analysis we explore the sensitivity of estimates to the choice of control group.

Using the group whose parents separate at grade 10 (perhaps augmented by those never treated) as control group, we can estimate the effect of experiencing separation in grade 4 on test scores in grades 4, 6 and 8; for those whose parents separate in grade 6, we can estimate effects in grades 6 and 8; and when parents separate in grade 8, we can estimate effects in grade 8.

Let D_{it}^{e} be an indicator variable that is 1 if dissolution happened at grade $t-e$, that is e years prior to grade t, and 0 otherwise. Thus, e is event time, and $e = t - g$ where g denotes the grade of dissolution. An event study model for the dynamic effects of separation at a particular grade $g \in \{4,6,8\}$, using $g = 10$ as control group, can be written

$$Y_{it} = \alpha_i + \lambda_t + \sum_{e \in K(g)} \gamma_e D_{it}^{e} + X_{it}\beta + \varepsilon_{it} \tag{1}$$

where Y_{it} is the test score for individual i in grade t, α_i are individual fixed effects, λ_t are grade fixed effects, X_{it} is a vector of controls (in our application, fixed effects for the calendar year of the test), and ε_{it} is the error term. The parameters of interest are γ_e which can be interpreted as the average treatment effect on the treated (ATT) at event time e, given the common trend assumption. This assumption cannot be tested, but our investigation of pre-trends (based on estimates of γ_e for negative values of e) indicates that it is likely to hold.

We choose $e = -2$ (the grade just before separation) as reference category (leaving out $\gamma_{-2} D_{it}^{-2}$ from the model). Test scores are observed at grade $t \in \{2,4,6,8\}$. Therefore, if $g = 4$ (dissolution happened at grade 4), we have observations for one pre-treatment period and three post-treatment periods, and we can estimate effects on test scores for event times $e \in K(4) = \{0,2,4\}$ (relative to

$e = -2$). Similarly, when $g = 6$ (8) we have observations for 2 (3) pre-treatment and 2 (1) post-treatment periods, and $K(6) = \{-4,0,2\}$, and $K(8) = \{-6,-4,0\}$. Estimates of γ_e at $e = -4$ and $e = -6$ close to zero will indicate common pre-trends.

Thus, we can estimate three models like Eq. (1), restricted to the subsamples consisting of the treatment group defined by the grade of separation (4, 6 or 8) and the control group of those treated after grade 8 ($g = 10$). Equivalently, we can stack three datasets consisting of observations with ($g = 4$ or $g = 10$), ($g = 6$ or $g = 10$) and ($g = 8$ or $g = 10$), and define an identifier, r, taking the values 4, 6 and 8, for each dataset, and then estimate the same parameters in a single stacked regression of the form

$$Y_{itr} = \alpha_{ir} + \lambda_{tr} + \sum_{g \in \{4,6,8\}} \sum_{e \in K(g)} \gamma_{ge} D_{it}^{e} + X_{itr}\beta_r + \varepsilon_{itr} \tag{2}$$

Application of the Callaway and Sant'Anna estimation procedure

Our estimation strategy discussed above is flexible by allowing the treatment effects to vary by grade of treatment and by time since treatment. For comparison, we also estimate effects of parental separation using the method proposed by Callaway and Sant'Anna (2021) which is equally flexible. In addition, this method allows pre-treatment covariates to affect the outcome trends. In our application, this might be important if, for instance, test score trends (in the non-treatment state) differ by the level of parental education and the shares of high and low educated parents differ between the treatment and control group. For each group (in our application defined by the grade of parental separation, g) and each time period (in our application the grade of the test, t) we estimate a "group-time average treatment effect", $ATT(g,t)$, that is, the average treatment effect for group g at time t, using as control group all not-yet-treated individuals at time t. Each $ATT(g,t)$

parameter is estimated in two steps. First, we estimate the relation between the potential outcome trend (from grade $g-2$ to t) in the non-treated state and covariates using only observations for the control group of not-yet-treated by time t,

$$Y_{it} - Y_{i,g-2} = X_i\theta^{gt} + \epsilon_{it}^{gt} \tag{3}$$

where X_i are covariates determined prior to time g (including a constant term). In the second step, we calculate the predicted trend (conditional on covariates) for the treatment group g in the non-treated state as $X_i\hat{\theta}^{gt}$ using the parameter estimate from the regression (3) and observations of X_i for the treatment group, and then we estimate the $ATT(g,t)$ as the average of $Y_{it} - Y_{i,g-2} - X_i\hat{\theta}^{gt}$ using only observations for the treatment group g.[9]

The estimation of $ATT(g,t)$ discussed above presumes that $t \geq g$. However, similar methods are used to estimate differences in pre-trends between treatment and control groups for $t < g$. Instead of using $g-2$ as base, these parameters are in the Callaway and Sant'Anna approach estimated with $t-2$ as base. For instance, for treatment group $g = 8$ we estimate the (differential) pre-trends from $t = 2$ to $t = 4$ and from $t = 4$ to $t = 6$. Again, all not-yet-treated observations are used as control group implying that the first of these pre-trends, $ATT(8,4)$, is estimated using the groups $g = 6$ and $g = 10$ as control group, whereas the second, $ATT(8,6)$, is estimated using $g = 10$ as control group. These two pre-trend parameters represent estimates for event times -4 and -2. Note that the pre-trends in this model have a different interpretation than the pre-trends in the two-way fixed effects models (1) and (2) where $g-2$ is the base also for the pre-trends. The

[9] We focus here on the outcome regression approach. Callaway and Sant'Anna (2021) also consider inverse probability weighting and double-robust estimators. Using these alternative estimators produce very similar results in our application.

Callaway and Sant'Anna (2021) method incorporates a bootstrap procedure to account for the dependency across the estimators of the different treatment effects and the two-stage estimation of each treatment effect.

The $ATT(g,t)$ estimates can be aggregated in different ways. For instance, by event time calculated as averages over different treatment groups, by grade of separation calculated as averages over different event times, by grade of test calculated as averages over different treatment groups, and a single average effect over all dimensions. As suggested by Callaway and Sant'Anna (2021), we calculate these aggregate parameters using weighted averages with weights determined by treatment group size.

Counterfactual distribution of test scores absent union dissolution

We can further extend our understanding of the effect of union dissolution on test scores by investigating whether the effect differs by the level of test scores, that is, by child ability. Based on Chernozhukov et al. (2013) we estimate a series of regressions across the test score distribution, where, for a finite set of points, we predict how union dissolution affects the probability of having a test score below each point.[10] Based on Eq. (1), we can, for each $g \in \{4,6,8\}$ (with $g = 10$ as control group), estimate a series of regressions:

$$q_{itj} = \alpha_{ij} + \lambda_{tj} + \sum_{e \in K(g)} \gamma_{ej} D_{it}^{e} + X_{it}\beta_j + \varepsilon_{itj} \tag{4}$$

where q_{itj} are indicator variables, $q_{itj} = 1(Y_{it} \leq j)$ with j varying within the interval from $\min(Y_{it})$ to $\max(Y_{it})$. Based on these estimates, we can, across the test score distribution, predict the

[10] For other applications of the distribution regression approach, see, e.g., Duflo (2001), Almond et al. (2011), and Biewen et al. (2022). Callaway and Li (Callaway and Li 2019) discuss quantile treatment effects in DID models.

probability of a child in the treatment group scoring less than j with and without having experienced parental union dissolution. From (4) we can, for each value of g and e predict $q_j^1 = E\left(q_{itj} \middle| D_{it}^e = 1, \alpha_{ij}, \lambda_{tj}, X_{it}\right)$ and the counterfactual $q_j^0 = E\left(q_{itj} \middle| \gamma_{ej} = 0, D_{it}^e = 1, \alpha_{ij}, \lambda_{tj}, X_{it}\right)$. Under an assumption of strict exogeneity, we can calculate q_j^1 and q_j^0 over j (given g and e) and, assuming rank stability, obtain the cumulative density function of test scores for the treated $\left(q_j^1\right)$ and the counterfactual function for the treated had they not experienced parental union dissolution $\left(q_j^0\right)$. The difference between q_j^1 and q_j^0 is simply γ_{ej}. Thus, the estimates of γ_{ej} across j provide an indication of what part of the test score distribution contribute most to the overall estimate of the effect of separation on test scores.

Estimating the Immediate Effect of Separation: A Regression Discontinuity Approach

To estimate the immediate effect of parental separation, we utilize the fact that we have information on the precise dates of children's tests and also very precise information on the time of parental separation. Under the assumption that parents do not manipulate the timing of separation, it is possible to examine whether any impact on children's test scores occurs at the time of separation. If the separation event causes a shock to children's cognitive ability or their ability to perform in the test situation, this should show up as a discrete change in test scores occurring at the time of separation. We define an underlying running variable R_i that captures time from parental separation to time for sitting the test:

$$R_i = Date\ of\ test_i - Date\ of\ parental\ separation_i$$

Thus, $R_i = \bar{r} = 0$ if the test occurs the same date as the parents separate, $R_i < \bar{r}$ if the test occurs before the separation, and $R_i > \bar{r}$ if the test occurs after the separation. Thus, treatment occurs at $R_i \geq \bar{r}$, and the standard RDD treatment effect at the cutoff is:

$$\tau = \tau(\bar{r}) = E\{Y_i(1) - Y_i(0)|R_i = \bar{r}\}$$

where $Y_i(0)$ and $Y_i(1)$ are the potential outcomes for each child *i* had the parents not separated yet (0) or separated (1).

Identification relies on parents' not sorting into separation around the time of children sitting the test, either as a response to the test or due to other reasons coinciding with the test. As discussed in the Introduction, the tests are low stake so it is unlikely that parents should strategically time their separation relative to the test date. Also, results from applying McCrary's (2008) test do not indicate manipulation around the cutoff, and covariates are balanced. The discontinuity we investigate is rather noisy (fuzzy) which is not surprising given the heterogeneity in separation processes related to differences in causes and consequences of separation discussed above. We use automatic bandwidth selection and control for grade of test and grade of separation.

5. DATA AND SAMPLE

We base our paper on Danish data and combine two data sources. Firstly, we include data on child school performance from the Danish national tests in all public schools in the school years 2009/2010 to 2017/2018. The national tests were introduced in 2009 for all public schools (with an option to opt-in for private schools[11]) and aimed to provide a uniform measure of child school performance in Denmark. The tests are conducted in the spring season and consist of a battery of different tests varying over grades. We focus on Danish reading comprehension tests, which are conducted in grades 2, 4, 6, and 8. Tests are mandatory for all children enrolled in Danish public schools, are adaptive, and are conducted on computers with an automatically generated result (Beuchert and Nandrup 2014). The test results are not displayed in any formal school diplomas and

[11] Clearly distinguishing between public and private schools are not possible in the data.

have no direct implications for the child's further school opportunities (there is no tracking or ability grouping in Danish primary or lower secondary school). For these reasons, parents might be less likely to try to affect the result by investments such as tutoring, and the lack of parental attention or interference might make the national tests an accurate measure of child ability.[12] We standardize the test scores to mean zero and standard deviation one within grade and year.

Secondly, we link each child to Danish full population register data via personal identification numbers to identify parents to each child. Due to the unique feature of the data, we have complete linkage for the population We construct variables for parental education and income measured at child age 5 and, importantly, we use register data to identify the date of parental separation of both married and non-married couples with children. We define the date of separation as the date one parent registers a new address different from the address of the other parent and condition on parents living apart for the subsequent two years. Potentially, there is a misreporting problem in the exact date of separation. Parents might move apart before one parent is able to register a new address. In Denmark, people are mandated by law to report changing their address within five days of a move. Whereas the law is not heavily enforced, reporting is done easily online and failing to do so triggers a fine of 1000 DKK (US$ 170) and risking not receiving official mail related to e.g., child custody arrangements or child payments. Extensive misreporting would bias our estimated effects towards zero.

Sample and descriptive statistics

We restrict our sample to children for whom we can identify both parents, and we leave out children who repeat or skip one or more grades. Our total sample consists of 698,414 children taking

[12] Post-test, parents may react to test results (Andersen and Nielsen 2020).

1,562,919 tests. In the sample, 35% of the children experience a union dissolution, the majority before school starting age. As discussed in Section 4, in the DID analysis we categorize treatment status by whether separation occurred before grade 2, or at grade 2, 4, 6, 8 or 10, respectively, or whether parents were not separated within our sample period. Table 1 shows means of test scores and covariates by treatment status. Average test scores at grades 2-8 are lower for children experiencing separation earlier, and they are higher for those experiencing separation after grade 8 or never. Parental education and income (at child age 5) tend to be higher for those experiencing separation later or never. There is no clear trend in the variation in being young or old for grade across treatment status. These indicators are effectively determined by school entry in grade 0 (at age 5).

6. RESULTS

Long-term Effect: Results from Event Study Design

First, we consider more long-term consequences of a parental union dissolution. We estimate event study regressions with individual fixed effects, where we assume common trends between treatment groups experiencing a parental union dissolution at a given grade and control groups experiencing this at later grades (and in some estimations we also include never treated in the control group). As discussed above in Section 4, we only include children who we observe in the data prior to experiencing a union dissolution. Thus, we can estimate effects of separation at grades 4, 6 and 8.

Stacked fixed effects regressions

Table 2 reports estimates of effects of separation on test scores by event time and grade of separation from stacked regressions as Equation (2). The first column shows results when the control group consists of children experiencing parental divorce at grade 10, and in the second

column the control group also includes the large group of never treated.[13] The estimates in the first three rows relate to differential pre-trends between the treatment and control groups. Each of these estimates are insignificant in the first column, and we cannot reject a hypothesis that they are all zero, which supports the common trend assumption when the control group consists of later treated students. However, when the control group includes in addition the large group of never treated (column 2), the coefficient at event time -6 (based on the treatment group $g = 8$) is large and clearly significant. Together with the small and insignificant coefficients at event time -4, this indicates a negative trend in test scores from grade 2 to grade 4 (but no significant trend from grade 4 to grade 6) for those experiencing separation at grade 8 (compared to the control group). Thus, the common trend assumption seems less plausible when the control group includes the never treated.

For each treatment group defined by the grade of separation, we can estimate the effect at event time 0. The point estimate for $e = 0$ and $g = 4$ in the first column indicates that separation at grade 4 reduces test scores in grade 4 by 0.0425 SD, whereas the point estimates for $e = 0$ and $g = 6$ or $g = 8$ tend to be smaller indicating that the negative effects of separation are larger when separation occurs earlier, although these three estimates for event time zero are not statistically different (see the p-value for the "equality for e=0" test at the bottom of Table 2). The dynamic effects for treatment groups $g = 4$ and $g = 6$ indicate that the negative effects increase over time (the point estimates tend to increase in event time), and for the last group we can reject equality of coefficients for event times 0 and 2 (see the p-value in second to last row of Table 2). These patterns are similar for the model with the larger control group in column 2.

[13] We do not show results based on a control consisting only of never treated. These results are very similar to those in column 2 because the group of never treated is large (see Table 1).

For both models in Table 2, we cannot reject hypotheses of equal coefficients across treatment groups given event time (see the p-values for the three tests for event times -4, 0 and 2 at the bottom of the table). However, we do reject equality of all treatment effect estimates (for $e \geq 0$) and equality of effects across event time for treatment group 6 (and for treatment group 4 in column 2); see the bottom of Table 2.

Table 3 shows results for more simple models which impose equality of effects given event time. For both models in Table 3, the negative treatment effects increase by time since separation (from about -0.03 SD to about -0.06 SD), and this dynamic pattern is statistically significant (see the p-values at the bottom of the table). For each event time 0, 2 and 4, the size of the estimates is very similar across the two models. However, we still have a significant differential pre-trend parameter at event time -6 in the model including the never treated in the control group.

The models estimated in Tables 2 and 3 use the same control group for the estimation of all treatment effect parameters – the control group consists of those treated at grade 10, or it consists of those treated at grade 10 combined with the never treated. However, for the estimation of some of the treatment effect parameters, we may choose other control groups. For instance, when estimating the effect of separation at grade 4 on test scores in grade 4, we might use all not-yet treated as control group, that is, those treated in grades 6, 8 and 10. This would be in accordance with the suggestions in Callaway and Sant'Anna (2021). An alternative might be to only use those treated at grade 6 because this control group may be expected to be more similar to the control group in terms of observables and unobservables, making the common trend assumption more likely to hold. In Appendix Table A1 we show estimation results for all possible choices of control groups of not-yet treated for each parameter. The results are robust to the choice of control group. If we compare with the estimates using $g = 10$ as control group (see column 1 in Table 2, and columns 4, 7 and 10 in

Table A1), the alternative control groups tend to reduce the estimated effects at event time 0 and 2 for treatment group $g = 4$, but to increase the estimate for event time 0 for treatment group 6. However, these differences are not statistically significant. Again, we find that pre-trends are not statistically different between treatment and control groups.[14]

Results applying the Callaway and Sant'Anna (2021) approach

In this subsection we report results using the approach for DID models with multiple time periods developed by Callaway and Sant'Anna (2021). As discussed above in Section 4, this method is characterized by allowing time trends to depend on pre-treatment covariates (corresponding to letting λ_t in Eq. (1) be a function of such covariates), by estimating separate models for each combination of treatment grade and test grade (g, t) using as control group all not-yet-treated individuals at time t (or a group of never treated), by conducting simultaneous inference for all estimated parameters using a bootstrap procedure, and by aggregating the estimates in different dimensions using treatment group sizes as weights.

Table 4 shows our estimation results applying the Callaway and Sant'Anna procedure using outcome regression to control for covariates.[15] In columns 1 and 2 the control groups consist of all not-yet treated (excluding the never treated); in columns 3 and 4 the control groups also include never treated. For comparison with our stacked regression results in Tables 2 and 3, the models in columns 1 and 3 of Table 4 only allow the counterfactual trends to depend on the calendar year of test. In columns 2 and 4 the common trend assumption is conditional on all covariates (calendar

[14] The lower part of Table A1 and the note to this table specify the relation to the control groups that we use to estimate each $ATT(g, t)$ parameter in our application of the Callaway and Sant'Anna (2021) estimation procedure.
[15] Using as an alternative the double-robust estimator suggested by Callaway and Sant'Anna produces almost identical results (not shown but available upon request). We use the csdid command for Stata.

year, parental education and income at child age 5, child gender, and indicators for being young or old for grade).

The first three rows of estimates in Table 4 show $ATT(g,t)$ estimates for treatment for group $g = 4$ and $t = 4$, $t = 6$, and $t = 8$. Each of these estimates is based on the "long difference" in individual test scores from the pre-treatment period $(g - 2)$ to t; this is indicated in the parentheses in the left column. For instance, when $g = 4$ and $t = 8$ the long difference is from grade 2 to grade 8, indicated by (2-8). The next sets of results are for treatment groups $g = 6$, $g = 8$ and $g = 10$. As in Table 2 we have (for each column) six ATT estimates for $t \geq g$ (i.e., $e \geq 0$), and they have the same interpretation as the estimates in Table 2. However, the estimates for differential pre-trends have a different interpretation in Table 4 compared to Table 2. In Table 2 the pre-trend parameters are measured relative to grade $g - 2$, whereas in Table 4 they are measured as a "short difference" relative to the previous pre-treatment period. For instance, for treatment group $g = 8$ we can estimate pre-trends from $t = 2$ to $t = 4$ (using as control group all not-yet treated at $t = 4$, i.e., those with $g = 6$ and $g = 10$ and possibly the never treated) and from $t = 4$ to $t = 6$ (using the control group $g = 10$ and possibly the never treated).

The results in column 1 of Table 4 are similar to those in column 1 of Table 2, although they are not identical because estimation methods and control groups differ. For instance, the $ATT(4,4)$ point estimate in Table 4 is -0.0325 while in Table 2 the corresponding estimate is -0.0425. This difference can mainly be explained by differences in control groups; in column 2 of Appendix Table A1 reporting results for a fixed-effects regression with a control group of all not-yet-treated, the estimate is -0.0322 and thus very close to the estimate in Table 4. Differences across event time tend to be smaller in column 1 of Table 4 compared to column 1 of Table 2. For instance, the estimate of $ATT(4,8)$ in Table 4 is -0.0552, whereas the corresponding estimate in Table 2 is

-0.0718. Allowing the trend to depend on more covariates (as in column 2 of Table 4) does not change estimates much. Including the large group of never treated in the control group does not have a large effect on estimates either.

For all four columns in Table 4, we cannot reject a hypothesis that all differential pre-trend parameters are zero (despite the significant estimates for $ATT(8,4)$ in columns 3 and 4); see the p-values reported at the bottom of Table 4.[16] As for the models in Table 2, we cannot reject equality of effects given event time either. For columns 1 and 2 we cannot reject equality of all ATT parameters for $e \geq 0$, whereas this hypothesis is clearly rejected when we enlarge the control group with the never treated in columns 3 and 4; see the p-values in the last row of Table 4. Thus, the results in columns 3 and 4 of Table 4 indicate (as the results in Table 2) that there are significantly increasing effects in event time.

Table 5 and Figure 1 show the estimates of Table 4 aggregated by event time. All estimates for $e \geq 0$ are clearly significant, and estimates are increasing in event time, although differences across event time are smaller when the group of never treated does not enter the control group. When the never treated are included in the control group, the differential pre-trend at event time -4 is rather large and significant, but the estimates at event times -6 and -2 are very small and clearly insignificant and, as discussed above, we cannot reject a hypothesis that all pre-trend parameters are zero. At the bottom of Table 5 we show the overall ATT estimates (for $e \geq 0$) obtained by aggregating the estimates of Table 4 across event time as well as by event time. When the control

[16] This conclusion is not changed if we instead test if the sum of the pre-trend parameters within each treatment group are zero, i.e., if we test the joint hypothesis: $ATT(10,8) + ATT(10,6) + ATT(10,4) = 0 \wedge ATT(8,6) + ATT(8,4) = 0 \wedge ATT(6,4) = 0$ (where the term $ATT(10,8)$ is omitted for the models in columns 1 and 2). This test of pre-trends is more similar to the test of pre-trends in the stacked regression DID models of Table 2. The p-values of this test in the models of Table 4 are: 0.564, 0.555, 0.200, and 0.289.

groups consist of later treated (columns 1 and 2), the overall effect is about -0.035 SD, and when it also includes never treated the overall effect is about -0.042 SD. This difference is mainly driven by differential effects at event times 2 and 4, while the effects at event time 0 are similar across models. Overall, the size of the estimates and their pattern by event time are very similar across specifications.

Distributional effects

To investigate at what levels of test scores parental union dissolution has important effects, we now present results of applying distribution DID regressions with later treated as control group; see the discussion in Section 4. Figure 2 shows on the y-axis the effect of parental union dissolution at grade 4 on the probability of scoring below the test score level indicated on the x-axis. The graphs in the first column of Figure 2 show effects at event time 0 using as control group those experiencing separation at grades 6, 8, and 10, respectively. The second column shows effects at event time 2 (control groups $g = 8$ and $g = 10$), and the third column shows effects at event time 4 (for control group $g = 10$). Across event time, and control group, the results indicate that it is the middle of the test score distribution that accounts for the negative effect of parental union dissolution on children's test scores. Given the outcome variable is normalized with mean 0 and standard deviation 1, it is roughly the middle 50-60% of the distribution that carries the effect, indicating that the test scores of both high and low ability pupils are on average unaffected by parents ending their union. Possible mechanisms may be that high-ability students are likely to be more resilient, and that low school performance may tend to be related to poor home environments that may on average not be worsened much by parental separation; see the discussion in Section 2.

Figure A1 in the Appendix shows that these conclusions hold if we extend the control group of the first two rows to include all later treated (in row 1 we use the control group $g \in \{6,8,10\}$, and in

row 2 the control group $g \in \{8,10\}$). Figures A2 and A3 in the Appendix show similar results when the treatment group consists of those experiencing separation at grades 6 and 8, respectively. For event times 0 and 2 the results are somewhat weaker, but again they indicate that it is mainly the middle part of the distribution that drives the effect. For negative event times (i.e., for differential pre-trends), there is no significant pattern in the estimated distributional effects.

Differential effects at event time zero

Table A2 in the Appendix show results for stacked DID regressions that correspond to those in Table 2 except that we allow effects at event time 0 to differ depending on whether separation occurred before or after the test date. This is partly to reconcile our DID results with our RDD finding of a discontinuity in test scores when separation date equals test date; we discuss the RDD results below. Point estimates in Table A2 indicate larger effects when separation occurred before the test, but for treatment groups $g = 4$ and $g = 6$, and for all three treatment groups together, the difference is not significant. For treatment group $g = 8$ the difference is marginally significant at the 10% level when the control group consists of later treated, and at the 5% level when it also includes never treated. Thus, these estimates are consistent with an RDD effect at the point where the separation date equals the test date. But at the same time, the results show that our simpler main specification where we assume constant effects at event time zero (irrespective of whether separation occurred before or after the test) is reasonable. At positive and negative event times, the results in Table A2 are almost identical to the results in our main analysis in Table 2.

As in our main analysis, we cannot reject equality of effects by event time. Appendix Table A3 shows results for this restricted model which is more in line with the RDD analysis where we pool observations across grade of separation. In the restricted DID model, we reject a hypothesis that the

effect at event time 0 is the same for those experiencing separation before the test as for those experiencing separation after the test, consistent with the RDD results.

Robustness check: anticipation effects

As discussed in Section 2, in some cases parental separation may be preceded by a period of escalation of social problems or conflict in the family. Therefore, child test scores could be affected negatively in this period before separation. In our DID analysis we effectively assume that this period is no longer than about a year. If it is in fact longer for many children, the pre-treatment test scores at event time -2 would be indirectly affected by separation and our ATT estimates would be biased towards zero. This would correspond to an "anticipation" effect of separation already at event time -2. As a robustness check we have therefore estimated the effect of separation at grade 6 using test scores in grade 2 (instead of grade 4) as the comparison. When the control group consists of those treated in grade 10, we can estimate the effect on test scores in grade 6, and when the control group consists of never treated, we can estimate the effect on test scores in grades 6 and 8 (i.e., at event time 0 and 2). The results are shown in Appendix Table A4. There are no significant differences compared to the corresponding main estimates in Table 2 indicating that periods of "anticipation" of more than a year are not important.[17]

Immediate Effect of Parental Separation: RDD Regressions

To study the effect of immediate parental separation on test scores, we rely on a an RDD analysis, where the running variable *R* is the difference between the date of sitting the test and the date when parents move apart. We control for grade of test and grade of separation. We use the rdrobust

[17] With those treated in grade 10 as controls, the point estimate (SE) in Table A4 is -0.030 (0.020). This point estimate is larger than the corresponding estimate in Table 2 of -0.021, but the estimates are not significantly different. With never treated as controls, the point estimates in Table A4 at event times 0 and 2 are -0.036 and -0.063, which are very close to the corresponding estimates in Table 2 (-0.036 and -0.059).

command of Calonico et al. (2017) with automatic bandwidth selection. Table 6 shows results for models with triangular or uniform kernel, and with polynomials of order 1 or 2 in the running variable, reporting both conventional and bias-corrected estimates. We find a significant effect at the threshold (where the date of separation is the same as the test date) of between -0.03 and -0.04 SD. The automatically selected bandwidth varies between 108 and 207 weeks.

Corresponding to the models in Table 6, Figure 3 plots the weekly-binned average test scores (adjusted for grade of test and grade of separation) across the running variable (R) with an imposed cutoff at the time when separation date equals test date, and with a fitted linear or quadratic curve on each side of the cutoff. As seen from the figure, there is evidence of a discrete decline in test scores at $R = 0$ of about 3% of a SD. The discontinuity is fuzzy which may be explained by the considerable heterogeneity in dissolution processes and by the fact that test scores are determined by many other factors than parental dissolution.

At the bottom of Table 6 we report p-values of the McCrary (2008) test for indications of parents manipulating the timing of separation relative to the date of the test. There is no statistically detectable evidence of manipulation. As a balancing test, Table A5 in the Appendix reports results from RDD regressions for parental and child characteristics. There is no evidence of a discontinuity in covariates occurring at the cutoff.[18] In total, we find some indications of an immediate decline in test scores following parental separation. The basic RDD estimates in the first row of Table 6 are numerically smaller than the long-term effects from the DID models, but still substantial.

[18] For maternal income the jump test is marginally significant at the 10% level. However, the size of the jump is small (410 € in yearly total income or less than 1% of the average of 45,000 €, see Table 1) and the sign of the jump is opposite to the sign the insignificant jump in paternal income.

7. CONCLUSION AND DISCUSSION

We find that parental separation and subsequent union dissolution has both immediate and long-term negative effects on children's test scores in the range 2.5-6.8 % of a standard deviation. Effect sizes are relatively large and driven by children in the middle of the test score distribution. Our estimation strategy using dynamic DID models, taking account of individual fixed effects, as well as RDD models, overcome important difficulties in estimating effects of parental separation. Thus, there exists not only selection into families that experience union dissolution, but also selection into timing of separation, and effects may be dynamic, which may lead to confounding issues in more traditional approaches to estimating the effect of union dissolution on children.

In this study, we have extensively demonstrated that parental union dissolution has substantial and likely dynamic negative effects on children's learning. Initial parameter estimates are of an absolute size similar to the effect of small classrooms in the STAR experiment (Krueger 1999), the effect of an increase in teacher-quality of 0.5 of a standard deviation (Rockoff 2004), or the average effect of offering high-dosage tutoring (Fryer and Howard-Noveck 2020), and the dynamic effects hint at even larger learning losses as time from union dissolution increase. More than 30 % of Danish children can expect their parents to dissolve their union before the children leave school. For most of these children, the parental break up occurs before they even start school, at a time where parental influence and investments likely are even more important for skill development (Cunha and Heckman 2007) than at the margin that we have considered in this study. Insofar as the initial and dynamic nature of the effect of parental union dissolution on children's learning extends backwards similarly, the learning loss that is occurring due to parental union dissolution are likely substantially higher than what we have identified.

The findings of this study raise the question of how to address the disparities in children's skill formation caused by parental union dissolution. Forcing parents to remain in a relationship is both at odds with central tenets of modern ethics and family policies but would also require the assumption that parents' relationship would not degenerate to a state as harmful for children's learning as it would be if the parents dissolved their union. Previous work studying the first modern divorce reforms has generally found no effect on educational attainment in the very first birth cohorts whose parents were affected by the reform change (Gruber 2004; Piketty 2003), perhaps because the divorcing marriages were bad, so children would be as well off with their parents being divorced (see also Clark et al. 2015). More recently and considering a setting where divorce is much less of a rare event and more normatively accepted, work from Denmark (Fallesen 2021), the Netherlands (Kabátek 2019), and South Korea (Lee 2013) all corroborate that forcing divorcing couples to have a mandatory separation or 'cooling-off' period causes about 10 percent of couples to reconsider their decision. If the home environment does not deteriorate further after such a change of heart, it might be a viable intervention to avoid "unnecessary" union dissolution, where parents change their minds if given time to reconsider.

This still leaves us with a substantial group of children of parents who would separate anyway. For this group, it is hard to envision a policy targeting their domestic situation that may help mitigate their loss of human capital due to their parents' union dissolution. This shifts the focus to what can be done in the school-environment. Teachers are often some of the first to know about changes to children's domestic situation. Increased intensity of schooling through remedial programs or scholastic focused after-school care could perhaps offset the loss to learning due to the parental dissolution. Non-academic investments improving child well-being might also have positive effects on academic achievement.

Lastly, we have shown that the children whose test scores decrease the most following their parents' union dissolution likely are the children found in the middle of the skill distribution. This is a group often ignored by educational policy and interventions, which instead often and with good reason considers how to increase performance among those in the lowest part of the distribution, and how to ensure that high performance students continue to be challenged and progress in the classroom. Yet, when it comes to parental union dissolution, it appears to be this middle group that experience most, if not all, of the negative consequences of parental union dissolution. This further means that for this group, the effect is likely much higher than the average effect that we estimate in our main analysis.

8. REFERENCES

Almond, Douglas, Hilary W. Hoynes, and Diane Whitmore Schanzenbach. 2011. "Inside the War on Poverty: The Impact of Food Stamps on Birth Outcomes." *The Review of Economics and Statistics* 93 (2): 387–403.

Alvord, Mary Karapetian, and Judy Johnson Grados. 2005. "Enhancing Resilience in Children: A Proactive Approach." *Professional Psychology: Research and Practice* 36 (3): 238–45. https://doi.org/10.1037/0735-7028.36.3.238.

Amato, Paul R. 1993. "Children's Adjustment to Divorce: Theories, Hypotheses, and Empirical Support." *Journal of Marriage and the Family* 55 (1): 23. https://doi.org/10.2307/352954.

———. 2000. "The Consequences of Divorce for Adults and Children." *Journal of Marriage and Family* 62 (4): 1269–87. https://doi.org/10.1111/j.1741-3737.2000.01269.x.

———. 2010. "Research on Divorce: Continuing Trends and New Developments." *Journal of Marriage and Family* 72 (3): 650–66. https://doi.org/10.1111/j.1741-3737.2010.00723.x.

Andersen, Simon Calmar, and Helena Skyt Nielsen. 2020. "Learning from Performance Information." *Journal of Public Administration Research and Theory* 30 (3): 415–31. https://doi.org/10.1093/jopart/muz036.

Bargain, Olivier, Libertad González, Claire Keane, and Berkay Özcan. 2012. "Female Labor Supply and Divorce: New Evidence from Ireland." *European Economic Review* 56 (8): 1675–91. https://doi.org/10.1016/j.euroecorev.2012.08.007.

Becker, Gary S., Elisabeth M. Landes, and Robert T. Michael. 1977. "An Economic Analysis of Marital Instability." *Journal of Political Economy* 85 (6): 1141–87. https://doi.org/10.1086/260631.

Beuchert, Louise Voldby, and Anne Brink Nandrup. 2014. "The Danish National Tests – A Practical Guide." *Economics Working Papers*, November. https://ideas.repec.org/p/aah/aarhec/2014-25.html.

Biewen, Martin, Bernd Fitzenberger, and Marian Rümmele. 2022. "Using Distribution Regression Difference-in-Differences to Evaluate the Effects of a Minimum Wage Introduction on the Distribution of Hourly Wages and Hours Worked." *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4219406.

Björklund, Anders, Donna K. Ginther, and Marianne Sundström. 2007. "Family Structure and Child Outcomes in the USA and Sweden." *Journal of Population Economics* 20 (1): 183–201. https://doi.org/10.1007/s00148-006-0094-7.

Björklund, Anders, and Marianne Sundström. 2006. "Parental Separation and Children's Educational Attainment: A Siblings Analysis on Swedish Register Data." *Economica* 73 (292): 605–24. https://doi.org/10.1111/j.1468-0335.2006.00529.x.

Bonnet, Carole, Bertrand Garbinti, and Anne Solaz. 2021. "The Flip Side of Marital Specialization: The Gendered Effect of Divorce on Living Standards and Labor Supply." *Journal of Population Economics* 34 (2): 515–73. https://doi.org/10.1007/s00148-020-00786-2.

Brown, David W., Robert F. Anda, Henning Tiemeier, Vincent J. Felitti, Valerie J. Edwards, Janet B. Croft, and Wayne H. Giles. 2009. "Adverse Childhood Experiences and the Risk of Premature Mortality." *American Journal of Preventive Medicine* 37 (5): 389–96. https://doi.org/10.1016/J.AMEPRE.2009.06.021.

Brüderl, Josef, and Frank Kalter. 2001. "The Dissolution of Marriages: The Role of Information and Marital-specific Capital." *The Journal of Mathematical Sociology* 25 (4): 403–21. https://doi.org/10.1080/0022250X.2001.9990262.

Cáceres-Delpiano, Julio, and Eugenio Giolito. 2012. “The Impact of Unilateral Divorce on Crime.” *Journal of Labor Economics* 30 (1): 215–48. https://doi.org/10.1086/662137.

Callaway, Brantly, and Tong Li. 2019. “Quantile Treatment Effects in Difference in Differences Models with Panel Data.” *Quantitative Economics* 10 (4): 1579–1618. https://doi.org/10.3982/QE935.

Callaway, Brantly, and Pedro H.C. Sant’Anna. 2021. “Difference-in-Differences with Multiple Time Periods.” *Journal of Econometrics* 225 (2): 200–230. https://doi.org/10.1016/J.JECONOM.2020.12.001.

Calonico, Sebastian, Matias D Cattaneo, Ann Arbor, Max H Farrell, and Rocío Titiunik. 2017. “Rdrobust: Software for Regression-Discontinuity Designs.” *The Stata Journal* 17 (2): 372–404.

Cattaneo, Matias D., Michael Jansson, and Xinwei Ma. 2018. “Manipulation Testing Based on Density Discontinuity.” *The Stata Journal: Promoting Communications on Statistics and Stata* 18 (1): 234–61. https://doi.org/10.1177/1536867X1801800115.

Cederblad, Marianne, Lisa Dahlin, Olle Hagnell, and Kjell Hansson. 1995. “Intelligence and Temperament as Protective Factors for Mental Health. A Cross-Sectional and Prospective Epidemiological Study.” *European Archives of Psychiatry and Clinical Neuroscience* 245 (1): 11–19. https://doi.org/10.1007/BF02191539.

Chaisemartin, Clément de, and Xavier D’Haultfœuille. 2020. “Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects.” *American Economic Review* 110 (9): 2964–96. https://doi.org/10.1257/aer.20181169.

Chernozhukov, Victor, Iván Fernández-Val, and Blaise Melly. 2013. “Inference on Counterfactual Distributions.” *Econometrica* 81 (6): 2205–68. https://doi.org/10.3982/ECTA10582.

Clark, Andrew E, Warn Lekfuangfu, Nattavudh Powdthavee, and George Ward. 2015. “Breaking up for the Kids’ Sake: Evidence from a British Birth Cohort.” Working paper. Paris School of Economics.

Condly, Steven J. 2006. “Resilience in Children: A Review of Literature with Implications for Education.” *Urban Education* 41 (3): 211–36. https://doi.org/10.1177/0042085906287902.

Crouch, Elizabeth, Janice C. Probst, Elizabeth Radcliff, Kevin J. Bennett, and Selina Hunt McKinney. 2019. “Prevalence of Adverse Childhood Experiences (ACEs) among US Children.” *Child Abuse & Neglect* 92 (June): 209–18. https://doi.org/10.1016/J.CHIABU.2019.04.010.

Cunha, Flavio, and James Heckman. 2007. “The Technology of Skill Formation.” *American Economic Review* 97 (2): 31–47. https://doi.org/10.1257/AER.97.2.31.

Duflo, Esther. 2001. “Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment.” *The American Economic Review* 91 (4): 795–813.

Fallesen, Peter. 2021. “Who Reacts to Less Restrictive Divorce Laws?” *Journal of Marriage and Family* 83 (2): 608–19. https://doi.org/10.1111/jomf.12722.

Fallesen, Peter, and Richard Breen. 2016. “Temporary Life Changes and the Timing of Divorce.” *Demography* 53 (5): 1377–98. https://doi.org/10.1007/s13524-016-0498-2.

Fallesen, Peter, and Michael Gähler. 2020a. “Post-Divorce Dual-Household Living Arrangements and Adolescent Wellbeing.” In *European Studies of Population*, 337–54. Springer International Publishing. https://doi.org/10.1007/978-3-030-25838-2_16.

———. 2020b. “Family Type and Parents’ Time with Children: Longitudinal Evidence for Denmark.” *Acta Sociologica (United Kingdom)* 63 (4): 361–80. https://doi.org/10.1177/0001699319868522.

Felitti, Vincent J., Robert F. Anda, Dale Nordenberg, David F. Williamson, Alison M. Spitz, Valerie Edwards, Mary P. Koss, and James S. Marks. 1998. "Relationship of Childhood Abuse and Household Dysfunction to Many of the Leading Causes of Death in Adults: The Adverse Childhood Experiences (ACE) Study." *American Journal of Preventive Medicine* 14 (4): 245–58. https://doi.org/10.1016/S0749-3797(98)00017-8.

Friborg, Oddgeir, Dag Barlaug, Monica Martinussen, Jan H. Rosenvinge, and Odin Hjemdal. 2005. "Resilience in Relation to Personality and Intelligence." *International Journal of Methods in Psychiatric Research* 14 (1): 29–42. https://doi.org/10.1002/mpr.15.

Fryer, Roland G., and Meghan Howard-Noveck. 2020. "High-Dosage Tutoring and Reading Achievement: Evidence from New York City." *Journal of Labor Economics* 38 (2): 421–52. https://doi.org/10.1086/705882.

Garriga, Anna, and Fulvia Pennoni. 2022. "The Causal Effects of Parental Divorce and Parental Temporary Separation on Children's Cognitive Abilities and Psychological Well-Being According to Parental Relationship Quality." *Social Indicators Research* 161 (2): 963–87. https://doi.org/10.1007/s11205-020-02428-2.

Gerra, G., R. Caccavari, R. Delsignore, M. Passeri, G. Fertonani Affini, D. Maestri, C. Monica, and F. Brambilla. 1993. "Parental Divorce and Neuroendocrine Changes in Adolescents." *Acta Psychiatrica Scandinavica* 87 (5): 350–54. https://doi.org/10.1111/J.1600-0447.1993.TB03385.X.

González, Libertad, and Tarja K. Viitanen. 2009. "The Effect of Divorce Laws on Divorce Rates in Europe." *European Economic Review* 53 (2): 127–38. https://doi.org/10.1016/J.EUROECOREV.2008.05.005.

Goodman-Bacon, Andrew. 2021. "Difference-in-Differences with Variation in Treatment Timing." *Journal of Econometrics*, Themed Issue: Treatment Effect 1, 225 (2): 254–77. https://doi.org/10.1016/j.jeconom.2021.03.014.

Graaf, Paul M. de, and Matthijs Kalmijn. 2006. "Divorce Motives in a Period of Rising Divorce." *Journal of Family Issues* 27 (4): 483–505. https://doi.org/10.1177/0192513X05283982.

Gruber, Jonathan. 2004. "Is Making Divorce Easier Bad for Children? The Long-Run Implications of Unilateral Divorce." *Journal of Labor Economics* 22 (4): 799–833. https://doi.org/10.1086/423155.

Härkönen, Juho, Fabrizio Bernardi, and Diederik Boertien. 2017. "Family Dynamics and Child Outcomes: An Overview of Research and Open Questions." *European Journal of Population*. Springer Netherlands. https://doi.org/10.1007/s10680-017-9424-6.

Heinesen, Eskil. 2019. "Parental Separation and Child Outcomes: Education, Employment and Juvenile Delinquency." 143.

Heuveline, Patrick, and Jeffrey M Timberlake. 2004. "The Role of Cohabitation in Family Formation: The United States in Comparative Perspective." *Journal of Marriage and the Family* 66 (5): 1214–30. https://doi.org/10.1111/j.0022-2445.2004.00088.x.

Johnson, John H., and Christopher J. Mazingo. 2000. "The Economic Consequences of Unilateral Divorce for Children." SSRN Scholarly Paper. Rochester, NY. https://doi.org/10.2139/ssrn.236227.

Kabátek, Jan. 2019. "Divorced in a Flash: The Effect of the Administrative Divorce Option on Marital Stability in the Netherlands." IZA Discussion Paper Series. IZA. www.iza.org.

Katz, L. F., and J. M. Gottman. 1997. "Buffering Children from Marital Conflict and Dissolution." *Journal of Clinical Child Psychology* 26 (2): 157–71. https://doi.org/10.1207/s15374424jccp2602_4.

Kennedy, Sheela, and Steven Ruggles. 2014. "Breaking Up Is Hard to Count: The Rise of Divorce in the United States, 1980–2010." *Demography* 51 (2): 587–98. https://doi.org/10.1007/s13524-013-0270-9.

Kim, Hyun Sik. 2011. "Consequences of Parental Divorce for Child Development." *American Sociological Review* 76 (3): 487–511. https://doi.org/10.1177/0003122411407748.

Kneip, Thorsten, Gerrit Bauer, and Steffen Reinhold. 2014. "Direct and Indirect Effects of Unilateral Divorce Law on Marital Stability." *Demography* 51 (6): 2103–26. https://doi.org/10.1007/s13524-014-0337-2.

Kreidl, Martin, Martina Štípková, and Barbora Hubatková. 2017. "Parental Separation and Children's Education in a Comparative Perspective: Does the Burden Disappear When Separation Is More Common?" *Demographic Research* 36: 73–110.

Krueger, Alan B. 1999. "Experimental Estimates of Education Production Functions." *The Quarterly Journal of Economics* 114 (2): 497–532.

Kvist, Anette Primdal, Helena Skyt Nielsen, and Marianne Simonsen. 2013. "The Importance of Children's ADHD for Parents' Relationship Stability and Labor Supply." *Social Science and Medicine* 88: 30–38. https://doi.org/10.1016/j.socscimed.2013.04.001.

Laird, Jessica, Nick Fabrin Nielsen, and Torben Heien Nielsen. 2020. "Differential Effects of the Timing of Divorce on Children's Outcomes: Evidence from Denmark." SSRN Scholarly Paper. Rochester, NY. https://doi.org/10.2139/ssrn.3567651.

Lee, Jungmin. 2013. "The Impact of a Mandatory Cooling-off Period on Divorce." *The Journal of Law and Economics* 56 (1): 227–43. https://doi.org/10.1086/667710.

Leopold, Thomas. 2018. "Gender Differences in the Consequences of Divorce: A Study of Multiple Outcomes." *Demography* 55 (3): 769–97. https://doi.org/10.1007/s13524-018-0667-6.

Lesner, Rune V. 2018. "The Long-Term Effect of Childhood Poverty." *Journal of Population Economics* 31 (3): 969–1004. https://doi.org/10.1007/s00148-017-0674-8.

Lew, Daphne, and Hong Xian. 2019. "Identifying Distinct Latent Classes of Adverse Childhood Experiences Among US Children and Their Relationship with Childhood Internalizing Disorders." *Child Psychiatry and Human Development* 50 (4): 668–80. https://doi.org/10.1007/S10578-019-00871-Y/TABLES/5.

Lyngstad, Torkild Hovde, and Marika Jalovaara. 2010. "A Review of the Antecedents of Union Dissolution." *Demographic Research* 23: 257–92. https://doi.org/10.4054/DemRes.2010.23.10.

Mallinson, David C, and Felix Elwert. 2022. "Estimating Sibling Spillover Effects with Unobserved Confounding Using Gain-Scores." *Annals of Epidemiology* 67: 73–80. https://doi.org/10.1016/j.annepidem.2021.12.010.

McCrary, Justin. 2008. "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test." *Journal of Econometrics* 142 (2): 698–714. https://doi.org/10.1016/j.jeconom.2007.05.005.

McKinnish, Terra G. 2004. "Occupation, Sex-Integration, and Divorce." *American Economic Review* 94 (2): 322–25. https://doi.org/10.1257/0002828041301911.

———. 2007. "Sexually Integrated Workplaces and Divorce." *Journal of Human Resources* XLII (2): 331–52. https://doi.org/10.3368/jhr.XLII.2.331.

McLanahan, Sara, Laura Tach, and Daniel Schneider. 2013. "The Causal Effects of Father Absence." *Annual Review of Sociology* 39 (July): 399–427. https://doi.org/10.1146/annurev-soc-071312-145704.

OECD. 2022. "OECD Family Database." 2022. https://www.oecd.org/els/family/database.htm.

Perelli-Harris, Brienna, and Nora Sánchez Gassen. 2012. “How Similar Are Cohabitation and Marriage? Legal Approaches to Cohabitation across Western Europe.” *Population and Development Review* 38 (3): 435–67. https://doi.org/10.1111/j.1728-4457.2012.00511.x.

Perelli-Harris, Brienna, Michaela Kreyenfeld, Wendy Sigle-Rushton, Renske Keizer, Trude Lappegård, Aiva Jasilioniene, Caroline Berghammer, and Paola Di Giulio. 2012. “Changes in Union Status during the Transition to Parenthood in Eleven European Countries, 1970s to Early 2000s.” *Population Studies*. https://doi.org/10.1080/00324728.2012.673004.

Piketty, Thomas. 2003. “The Impact of Divorce on School Performance: Evidence from France 1968-2002.” *Centre for Economic Policy Research, Discussion Paper Series*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=482863.

Raz-Yurovich, Liat. 2013. “Divorce Penalty or Divorce Premium? A Longitudinal Analysis of the Consequences of Divorce for Men’s and Women’s Economic Activity.” *European Sociological Review* 29 (2): 373–85. https://doi.org/10.1093/esr/jcr073.

Rockoff, Jonah E. 2004. “The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data.” *American Economic Review* 94 (2): 247–52. https://doi.org/10.1257/0002828041302244.

Rosenbeck, Bente. 2017. “Liberalization of Divorce: No-Fault Divorce in Denmark and the Nordic Countries in the Early 20th Century.” *Scandinavian Journal of History* 43 (1): 18–39. https://doi.org/10.1080/03468755.2017.1318338.

Sheppard, Paula, Justin R. Garcia, and Rebecca Sear. 2015. “Childhood Family Disruption and Adult Height: Is There a Mediating Role of Puberty?” *Evolution, Medicine and Public Health* 2015 (1): 332–42. https://doi.org/10.1093/emph/eov028.

Sigle-Rushton, Wendy, Torkild Hovde Lyngstad, Patrick Lie Andersen, and Oystein Kravdal. 2014. “Proceed with Caution? Parents’ Union Dissolution and Children’s Educational Achievement.” *Journal of Marriage and Family* 76 (1): 161–74. https://doi.org/10.1111/jomf.12075.

Smock, Pamela J., Wendy D. Manning, and Sanjiv Gupta. 1999. “The Effect of Marriage and Divorce on Women’s Economic Well-Being.” *American Sociological Review* 64 (6): 794. https://doi.org/10.2307/2657403.

Strohschein, Lisa. 2005. “Parental Divorce and Child Mental Health Trajectories.” *Journal of Marriage and Family* 67 (5): 1286–1300. https://doi.org/10.1111/J.1741-3737.2005.00217.X.

Sun, Y., and Y. Li. 2001. “Marital Disruption, Parental Investment, and Children’s Academic Achievement: A Prospective Analysis.” *Journal of Family Issues* 22 (1): 27–62. https://doi.org/10.1177/019251301022001002.

Svarer, Michael. 2007. “Working Late: Do Workplace Sex Ratios Affect Partnership Formation and Dissolution?” *Journal of Human Resources* XLII (3): 582–95. https://doi.org/10.3368/jhr.xlii.3.582.

Uunk, Wilfred. 2004. “The Economic Consequences of Divorce for Women in the European Union: The Impact of Welfare State Arrangements.” *European Journal of Population/ Revue Europenne de D*[illegible]*mographie* 20 (3): 251–85. https://doi.org/10.1007/s10680-004-1694-0.

Vezzetti, Vittorio Carlo. 2016. “New Approaches to Divorce with Children: A Problem of Public Health.” *Health Psychology Open* 3 (2): 2055102916678105. https://doi.org/10.1177/2055102916678105.

Wagner, Michael. 2020. “On Increasing Divorce Risks.” In *Divorce in Europe: New Insights in Trends, Causes and Consequences of Relation Break-Ups*, edited by Dimitri Mortelmans,

37–61. European Studies of Population. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-25838-2_3.

Weiss, Yoram, and Robert J. Willis. 1997. "Match Quality, New Information, and Marital Dissolution." *Journal of Labor Economics* 15 (1, Part 2): S293–329. https://doi.org/10.1086/209864.

Wolfers, Justin. 2006. "Did Unilateral Divorce Laws Raise Divorce Rates? A Reconciliation and New Results." *American Economic Review* 96 (5): 1802–20. https://doi.org/10.1257/aer.96.5.1802.

Zolkoski, Staci M., and Lyndal M. Bullock. 2012. "Resilience in Children and Youth: A Review." *Children and Youth Services Review* 34 (12): 2295–2303. https://doi.org/10.1016/J.CHILDYOUTH.2012.08.009.

Table 1. Means of test scores and covariates by grade of separation

	Time of Separation							
	Before grade 2	Grade 2	Grade 4	Grade 6	Grade 8	Grade 10	Never	Total
Test score	-0.115	-0.025	0.006	0.049	0.066	0.139	0.105	0.048
Boy	0.510	0.500	0.504	0.501	0.493	0.492	0.513	0.511
Mother no post-sec. degree	0.710	0.611	0.600	0.585	0.595	0.583	0.537	0.582
Father no post-sec. degree	0.760	0.685	0.679	0.663	0.661	0.640	0.613	0.651
Mother income	44.443	45.000	45.254	45.531	45.435	44.280	46.495	45.876
Father income	51.766	57.052	57.879	58.320	58.639	58.058	60.564	58.338
Young for grade	0.089	0.062	0.059	0.057	0.050	0.026	0.057	0.063
Old for grade	0.055	0.069	0.066	0.064	0.066	0.042	0.069	0.065
Observations	329,775	61,221	51,745	41,378	34,070	26,122	1,018,608	1,562,919
Children	148,507	27,241	22,939	17,627	13,850	11,439	456,811	698,414

Note. The table shows means of variables by grade of parental separation categorized using a “nearest test” criterion as discussed in Section 4. “Never” means that separation is not observed within our sample period. Test scores are standardized by grade and calendar year of test. The average test scores shown are averages over all test score observations; a child with four test score observations thus contributes with four observations to the average shown. Parental education and income are measured at age 5. Parental yearly total income (including public transfers) is deflated by the CPI (base 2015) and measured in €1,000 (exchange rate 7.44 DKK/€).

Table 2. Stacked regressions of event study models: Effects of separation on test scores by event time and grade of separation

Event time (e) / Grade of separation (g)	(1) Control group: Treated at grade 10	(2) Control group: Treated at grade 10 or never
-6 / 8	0.0160	0.0318***
	(0.0180)	(0.0108)
-4 / 6	0.00747	0.00443
	(0.0163)	(0.00856)
-4 / 8	-0.0102	0.00568
	(0.0118)	(0.00781)
-2	0	0
	(.)	(.)
0 / 4	-0.0425***	-0.0400***
	(0.0159)	(0.00802)
0 / 6	-0.0214*	-0.0355***
	(0.0116)	(0.00758)
0 / 8	-0.0264**	-0.0212**
	(0.0115)	(0.00835)
2 / 4	-0.0553***	-0.0628***
	(0.0176)	(0.00972)
2 / 6	-0.0540***	-0.0594***
	(0.0149)	(0.0105)
4 / 4	-0.0718***	-0.0662***
	(0.0217)	(0.0129)
N	160,435	2,785,063
Children	39,146	352,225
p-values of tests:		
All pre-trends 0	0.344	0.024
Equality for e = -4	0.410	0.915
Equality for e = 0	0.590	0.238
Equality for e = 2	0.952	0.810
Equality for e ≥ 0 at g = 4	0.231	0.012
Equality for e ≥ 0 at g = 6	0.012	0.017
Equality for e ≥ 0	0.044	0.001

Note. Results from stacked regressions of event study models controlling for individual fixed effects, grade fixed effects and calendar year of test fixed effects. The reference category for event time is -2 (the period just before onset of treatment). The number of observations and children is excluding singleton observations; the control group observations are included three times (one time for each treatment group). The last seven rows report p-values of tests on the pre-trend and treatment effect parameters. For instance, "all pre-trends 0" is a test that all three parameters for e<0 are equal to zero; "equality for e=0" is a test that all three parameters for event time 0 are equal; and "equality for e ≥ 0 at g = 6" is a test that the parameters at e=0 and e=2 for those first treated at grade 6 are equal. Standard errors in parentheses are clustered on students:

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3. Stacked regressions of event study models: Effects of separation on test scores by event time

Event time (e)	(1) Control group: Treated at grade 10	(2) Control group: Treated at grade 10 or never
-6	0.0180	0.0287***
	(0.0178)	(0.0102)
-4	-0.00441	0.00344
	(0.00818)	(0.00575)
-2	0	0
	(.)	(.)
0	-0.0291***	-0.0330***
	(0.00660)	(0.00459)
2	-0.0532***	-0.0590***
	(0.0115)	(0.00691)
4	-0.0664***	-0.0626***
	(0.0195)	(0.0120)
N	160,435	2,785,063
Children	39,146	352,225
p-values of tests:		
All pre-trends 0	0.197	0.017
Equality for e ≥ 0	0.010	0.000

Note. Results from stacked regressions of event study models controlling for individual fixed effects, grade fixed effects and calendar year of test fixed effects. The reference category for event time is -2 (the period just before onset of treatment). The number of observations and children is excluding singleton observations; the control group observations are included three times (one time for each treatment group). The last two rows report p-values of tests on the pre-trend and treatment effect parameters. Thus, "all pre-trends 0" is a test that the two parameters for e<0 are equal to zero; and "equality for e ≥ 0" is a test that the parameters at e=0, e=2 and e=4 are equal. Standard errors in parentheses are clustered on students:

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4. Callaway and Sant'Anna approach: Effects of parental separation on test scores by grade of separation and grade of test using as control groups later treated or later/never treated.

Treatment group (g) Grade of test (t)	Event time (e)	(1) Later	(2) Later	(3) Later/Never	(4) Later/Never
g=4					
t=4 (2-4)	0	-0.0325***	-0.0309***	-0.0398***	-0.0348***
		(0.0105)	(0.0105)	(0.0086)	(0.0084)
t=6 (2-6)	2	-0.0431***	-0.0402***	-0.0603***	-0.0570***
		(0.0131)	(0.0133)	(0.0102)	(0.0104)
t=8 (2-8)	4	-0.0552**	-0.0494**	-0.0698***	-0.0668***
		(0.0236)	(0.0240)	(0.0157)	(0.0155)
g=6					
t=4 (2-4)	-2	0.0124	0.0135	-0.0038	0.0006
		(0.0122)	(0.0122)	(0.0091)	(0.0092)
t=6 (4-6)	0	-0.0318***	-0.0313***	-0.0366***	-0.0364***
		(0.0095)	(0.0096)	(0.0079)	(0.0079)
t=8 (4-8)	2	-0.0491***	-0.0514***	-0.0586***	-0.0594***
		(0.0166)	(0.0168)	(0.0106)	(0.0109)
g=8					
t=4 (2-4)	-4	-0.0192	-0.0186	-0.0213**	-0.0178*
		(0.0125)	(0.0126)	(0.0098)	(0.0098)
t=6 (4-6)	-2	0.0070	0.0051	-0.0044	-0.0048
		(0.0123)	(0.0124)	(0.0078)	(0.0075)
t=8 (6-8)	0	-0.0175	-0.0191	-0.0136*	-0.0143*
		(0.0119)	(0.0120)	(0.0082)	(0.0083)
g=10					
t=4 (2-4)	-6	0.0165	0.0130	0.0004	0.0001
		(0.0161)	(0.0158)	(0.0140)	(0.0140)
t=6 (4-6)	-4	-0.0018	-0.0022	-0.0153	-0.0160
		(0.0124)	(0.0125)	(0.0098)	(0.0097)
t=8 (6-8)	-2	-	-	0.0052	0.0043
		(.)	(.)	(0.0078)	(0.0077)
Controls		Year	All	Year	All
p-values of tests:					
All pre-trends 0		0.441	0.511	0.232	0.345
Equality for e = -4		0.321	0.357	0.668	0.901
Equality for e = -2		0.756	0.631	0.634	0.693
Equality for e = 0		0.571	0.691	0.051	0.106
Equality for e = 2		0.776	0.603	0.909	0.876
Equality for e ≥ 0, g=4		0.622	0.724	0.138	0.096
Equality for e ≥ 0, g=6		0.366	0.300	0.096	0.089
Equality for e ≥ 0		0.535	0.644	0.001	0.002

Note. The table shows ATT(g,t) estimates using the method of Callaway and Sant'Anna (2021). In columns 1 and 2 the control group for each parameter consists of all later treated; in columns 3 and 4 it also includes never treated. In columns 1 and 3 the calendar year is the only control variable that may affect the counterfactual trend; in columns 2 and 4 all control variables are included: calendar year, parental education (4 categories) and income, child gender, and whether the child is young or old for grade. The numbers in the parentheses in the first column indicate the test score

difference estimated, for instance, (2-6) indicates the long difference from grade 2 to grade 6. The last rows show p-values for tests on the pre-trend and treatment effect parameters. Standard errors in parentheses are obtained using a Mammen wild bootstrap procedure (1,000 repetitions)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5. Callaway and Sant'Anna approach: Average effects by event time and overall, using as control groups later treated or later/never treated

Event time (e)	(1) Later	(2) Later	(3) Later/Never	(4) Later/Never
-6	0.0165	0.0130	0.0004	0.0001
	(0.0164)	(0.0160)	(0.0136)	(0.0135)
-4	-0.0110	-0.0109	-0.0185***	-0.0169**
	(0.0095)	(0.0095)	(0.0068)	(0.0068)
-2	0.0098	0.0094	-0.0012	-0.0001
	(0.0092)	(0.0091)	(0.0051)	(0.0051)
0	-0.0281***	-0.0277***	-0.0313***	-0.0295***
	(0.0058)	(0.0057)	(0.0049)	(0.0049)
2	-0.0458***	-0.0451***	-0.0595***	-0.0581***
	(0.0103)	(0.0104)	(0.0073)	(0.0074)
4	-0.0552**	-0.0494*	-0.0698***	-0.0668***
	(0.0251)	(0.0258)	(0.0140)	(0.0143)
Overall ATT	-0.0358***	-0.0348***	-0.0432***	-0.0413***
	(0.0071)	(0.0072)	(0.0050)	(0.0052)
Controls	Year	All	Year	All

Note. The table shows ATT estimates by event time using the method of Callaway and Sant'Anna (2021). The average effects by event time (and the overall average effect post treatment at the bottom of the table) are weighted averages of the relevant ATT(g,t) estimates in Table 4 with the size of the treatment groups as weights. In columns 1 and 2 the control group for each parameter consists of all later treated; in columns 3 and 4 it also includes never treated. In columns 1 and 3 the calendar year is the only control variable that may affect the counterfactual trend; in columns 2 and 4 all control variables are included: calendar year, parental education (4 categories) and income, child gender, and whether the child is young or old for grade. Standard errors in parentheses are obtained using a Mammen wild bootstrap procedure (1,000 repetitions)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

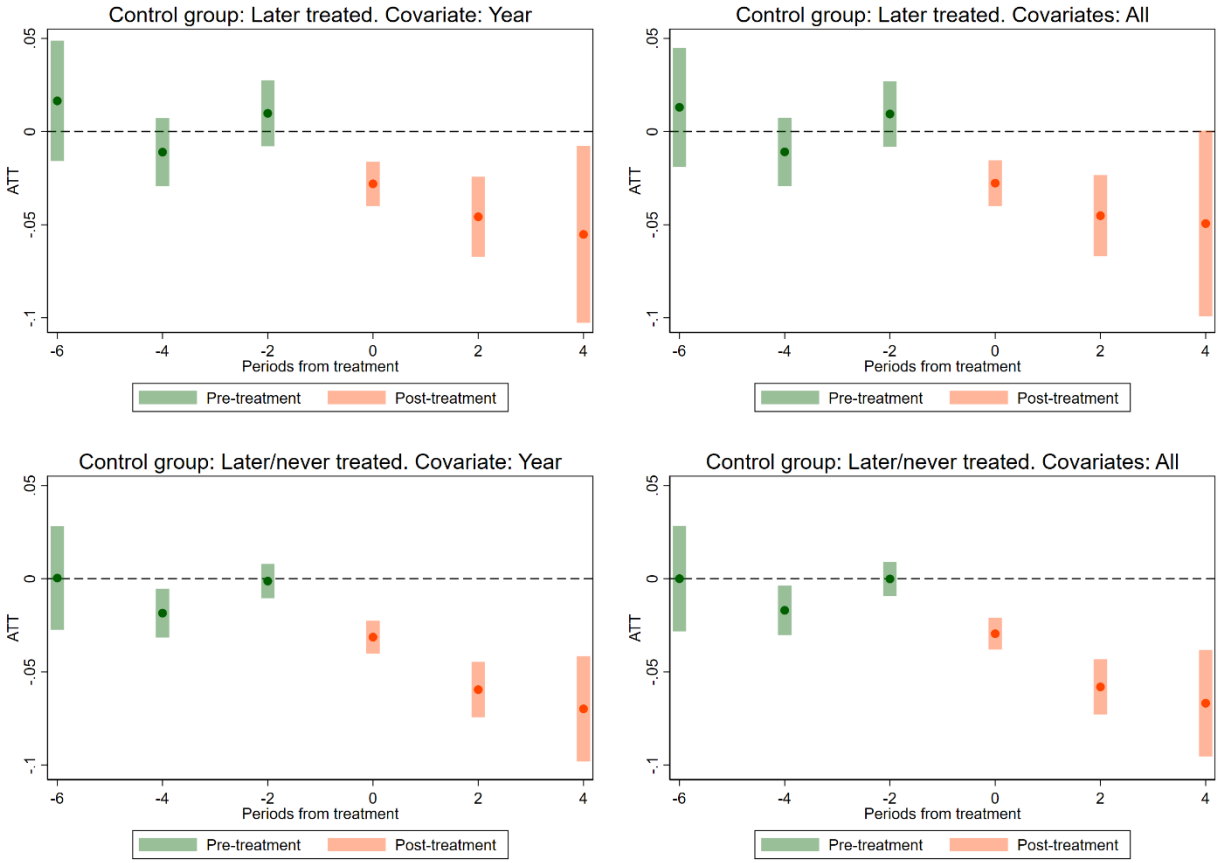
Table 6: Estimates from regression discontinuity design regressions: Coefficient on indicator for being above the threshold (date of test after date of separation)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Estimate	-0.029**	-0.029**	-0.038**	-0.034**	-0.033*	-0.037*	-0.040**	-0.049***
Standard error	(0.009)	(0.011)	(0.012)	(0.013)	(0.016)	(0.018)	(0.019)	(0.018)
Type	Conventional	Bias-corr	Conventional	Bias-corr	Conventional	Bias-corr	Conventional	Bias-corr
Kernel	Triangular		Uniform		Triangular		Uniform	
Polynomial	1		1		2		2	
Bandwidth, weeks	207.163		108.649		161.597		133.303	
McCrary test (p)	0.990		0.574		0.826		0.498	
Effective N	194,497		102,182		151,263		126,655	

Note: All regressions control for grade of test and nearest test relative to parental union dissolution. Estimated using rdrobust command by Calonico et al. (2017). Automatic bandwidth selection with mserd selector. McCrary test calculated using rddensity command by Cattaneo et al. (2018). Standard errors clustered at student level. Bias-corrected estimates reported with robust standard errors

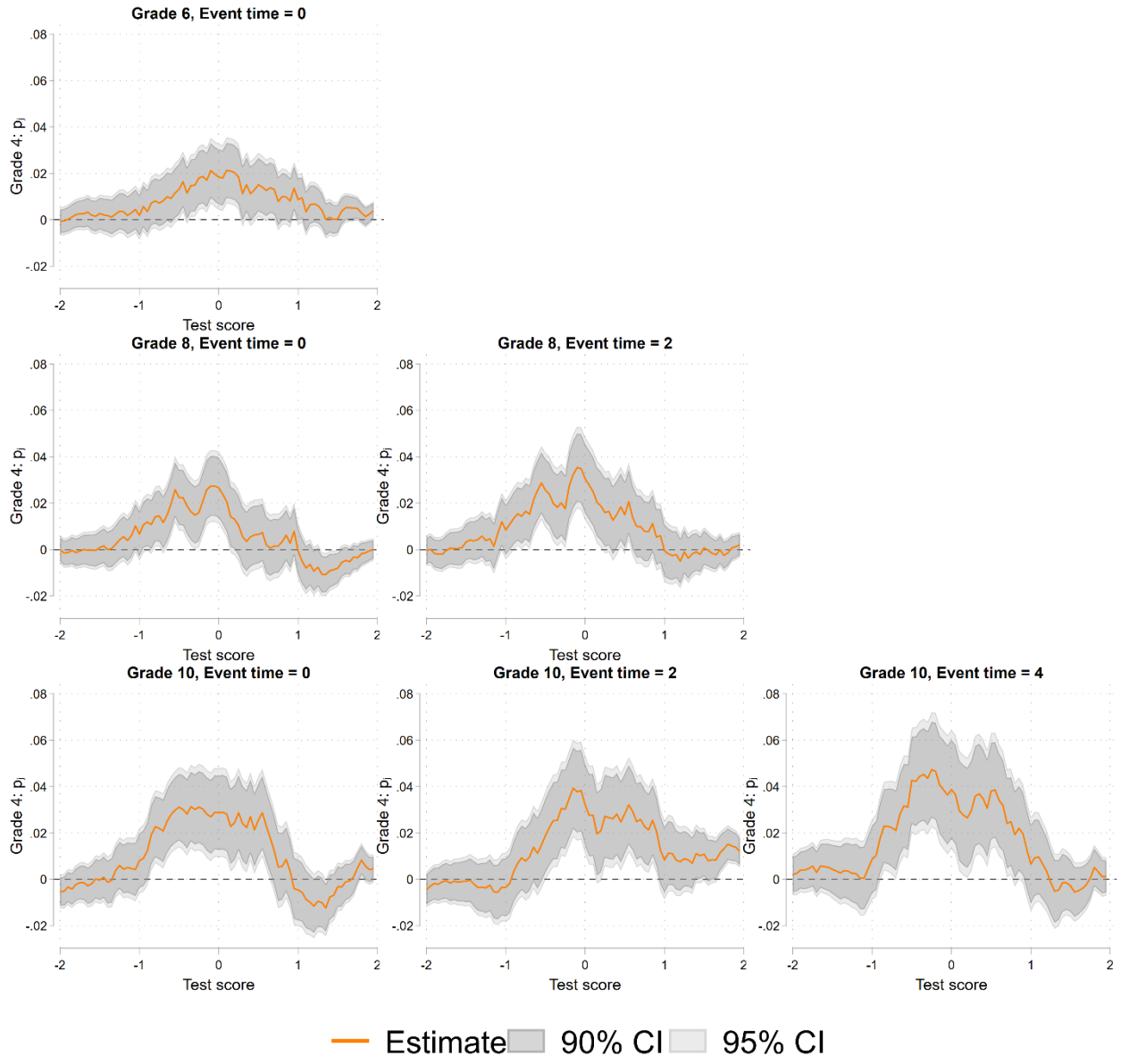
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure 1. Aggregate event time estimates by control group and pre-treatment covariates



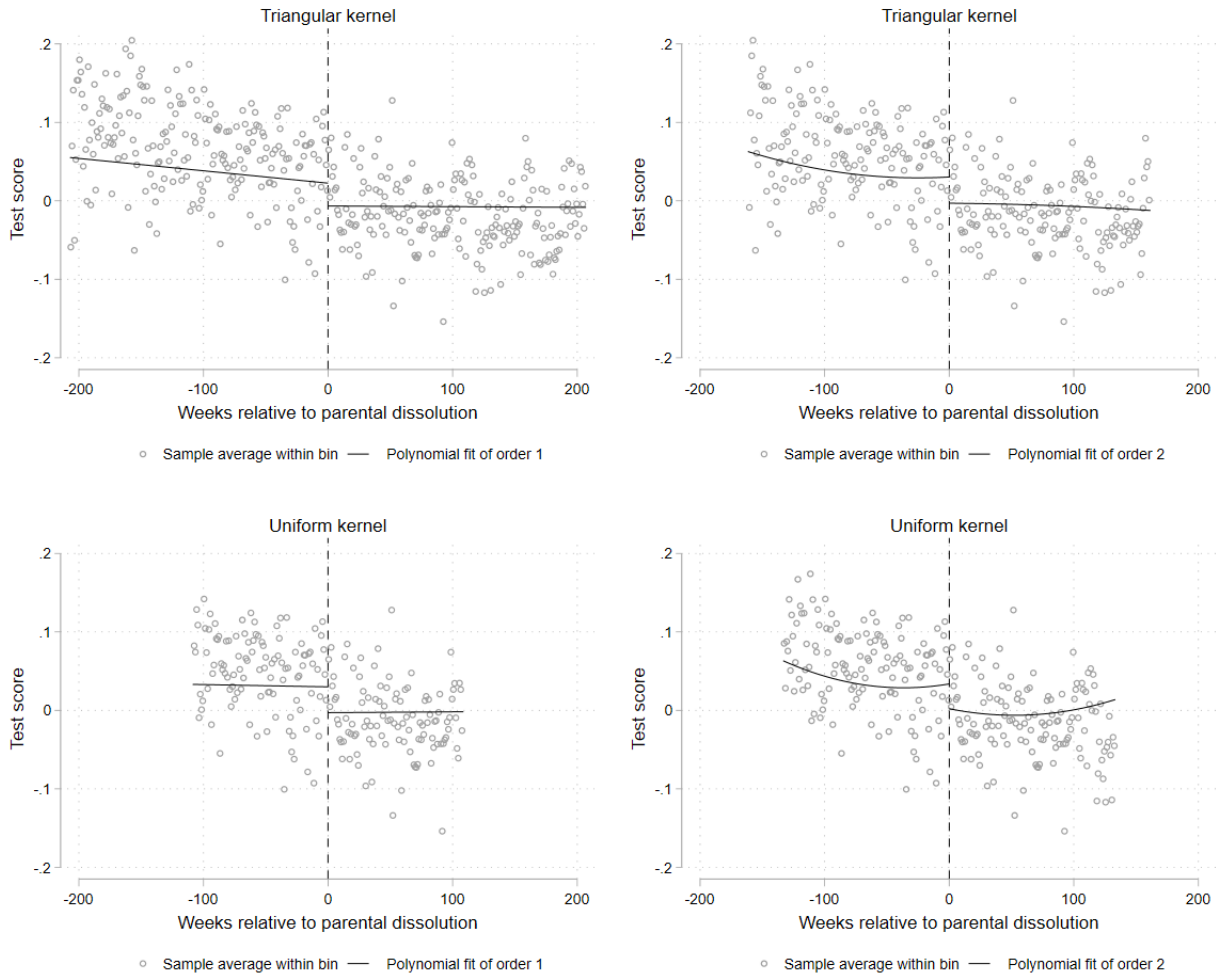
Note. This figure shows ATT estimates by event time for the four models of Table 5 using the method of Callaway and Sant'Anna (2021). The average effects by event time are averages of the ATT(g,t) estimates in Table 4 with the size of the treatment groups as weights. In the two upper graphs the control group for each parameter consists of all later treated; in the two lower graphs it also includes never treated. In the two graphs to the left the calendar year is the only control variable that may affect the counterfactual trend; in the two graphs to the right all control variables are included: calendar year, parental education (4 categories) and income, child gender, and whether the child is young or old for grade. The 95% confidence intervals are based on a Mammen wild bootstrap procedure (with 1,000 repetitions).

Figure 2. Effect of union dissolution in grade 4 on the probability of scoring below a given level across the distribution of test scores (comparison group treatment timing and event time in subplot title)



Note. The graphs show the effect of separation in grade 4 on the probability of scoring below the test score level indicated on the x-axis. Each subgraph shows effects for a given event time and a given comparison group consisting of later treated children. The graphs in the first column show effects at event time 0 using as control group those experiencing separation at grades 6, 8, and 10, respectively. The second column shows effects at event time 2 (control groups $g = 8$ and $g = 10$), and the third column shows effects at event time 4 (for control group $g = 10$). Confidence intervals based on robust parametric standard errors.

Figure 3. Plot of weekly-binned test scores and regression discontinuity lines under different kernels and polynomials



Note. Observations binned by week. All plots adjusted for grade of test and grade closest to parental union dissolution. Number of weeks displays estimated bandwidths.

9. APPENDIX

Table A1. Effects of separation on test scores by event time and grade of separation using varying control groups of later treated students

		(1)	(2)	(3)	(4)	(5)	(6)		(7)	(8)	(9)		(10)
g (treated)	4	4	4	4	4	4	4	6	6	6	6	8	8
g (control)		10	6,8,10	6	8	6,8	8,10		10	8,10	8		10
e	t							t				t	
-6												2	0.0160
													(0.0180)
-4								2	0.00747	-0.00779	-0.0163	4	-0.0102
									(0.0163)	(0.0118)	(0.0130)		(0.0118)
-2 (ref.)	2	0	0	0	0	0	0	4	0	0	0	6	0
		(.)	(.)	(.)	(.)	(.)	(.)		(.)	(.)	(.)		(.)
0	4	-0.0425***	-0.0322***	-0.0354***	-0.0198	-0.0297***	-0.0283**	6	-0.0214*	-0.0278***	-0.0301***	8	-0.0264**
		(0.0159)	(0.00989)	(0.0116)	(0.0126)	(0.0102)	(0.0113)		(0.0116)	(0.00951)	(0.0108)		(0.0115)
2	6	-0.0553***			-0.0387***		-0.0462***	8	-0.0540***				
		(0.0176)			(0.0145)		(0.0132)		(0.0149)				
4	8	-0.0718***											
		(0.0217)											
N (obs.)		52,407	52,194	35,148	47,414	46,568	60,576		54,630	61,153	47,991		53,398
N (students)		17,721	26,097	17,574	18,000	23,284	23,261		18,702	24,171	18,910		18,495
C&SA (g,t)		4,8	4,4				4,6		6,8	6,6			8,8
										6,4			8,6
													8,4

Note. Results from linear regressions of models controlling for individual fixed effects, grade fixed effects and calendar year of test fixed effects using varying control groups of later treated students. The reference category for event time is -2 (the period just before onset of treatment). In models (1)-(6) the treatment group consists of children experiencing parental separation in grade 4 (g=4). Models (7)-(9) are for the g=6 treatment group, and model (10) is for the g=8 treatment group. For each of the treatment groups we include a column showing the grade of test (t) corresponding to the event time. The estimates using g=10 as control group in models (1), (7) and (10) are identical to the stacked regression estimates presented in the first column of Table 2. The number of observations and children is excluding singleton observations. The last three rows indicate if the combination of treatment and control group in the column is used to estimate parameters in the model based on the Callaway and Sant'Anna method discussed in Section 4, and (if so) which ATT(g,t) parameters. In our application of this method, we also use the treatment group g=10 to estimate the pre-trend parameters ATT(10,6) using g=8 as control group, and ATT(10,4) using g=6 and g=8 as control group. Standard errors in parentheses are clustered on students: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A2. Stacked regressions: Effects of separation on test scores by event time and grade of separation. Effects at event time 0 are split into effects at event time -1 (separation occurred within a year after the test date) and event time 0 (separation occurred within a year prior to the test date)

Event time (e) / Grade of separation (g)	(1) Control group: Treated at grade 10	(2) Control group: Treated at grade 10 or never
-6 / 8	0.0160	0.0319***
	(0.0180)	(0.0108)
-4 / 6	0.00743	0.00442
	(0.0163)	(0.00856)
-4 / 8	-0.0103	0.00564
	(0.0118)	(0.00781)
-2	0	0
	(.)	(.)
-1 / 4	-0.0347*	-0.0378***
	(0.0187)	(0.0102)
-1 / 6	-0.0134	-0.0263**
	(0.0152)	(0.0105)
-1 / 8	-0.00645	-0.00497
	(0.0162)	(0.0118)
0 / 4	-0.0501***	-0.0421***
	(0.0191)	(0.0105)
0 / 6	-0.0286*	-0.0439***
	(0.0148)	(0.00993)
0 / 8	-0.0460***	-0.0371***
	(0.0161)	(0.0116)
2 / 4	-0.0553***	-0.0629***
	(0.0176)	(0.00972)
2 / 6	-0.0540***	-0.0595***
	(0.0149)	(0.0105)
4 / 4	-0.0718***	-0.0662***
	(0.0217)	(0.0129)
N	160,435	2,785,063
N_clust	39,146	352,225
p-values of tests:		
All pre-trends 0	0.340	0.024
Equality for e = -4	0.410	0.916
Equality for e ∈{-1,0} at g=4	0.453	0.743
Equality for e ∈{-1,0} at g=6	0.424	0.196
Equality for e ∈{-1,0} at g=8	0.080	0.049
Equality for e = -1	0.532	0.108
Equality for e = 0	0.654	0.903
Equality for e ∈{-1,0}	0.176	0.129
Equality for e = 2	0.953	0.812
Equality for e ≥ 0 at g = 4	0.478	0.108
Equality for e ≥ 0 at g = 6	0.110	0.192

Equality for e ≥ 0	0.501	0.164
Equality for e ≥ -1	0.022	0.001

Note. Results from stacked regressions of event study models controlling for individual fixed effects, grade fixed effects and calendar year of test fixed effects. The reference category for event time is -2. Compared to the models in Table 2 of the main analysis, we allow the effects at event time 0 to vary by whether separation occurred before or after the test. Thus, event time 0 in the main analysis represent observations where separation occurs from about 1 year prior to the test to about 1 year after the test. In the models of this appendix table, these observations are split up into event time -1 if separation occurs after the test, and event time 0 if separation occurs before the test (or at the same date). Otherwise, the specification is the same as in Table 2. The number of observations and children is excluding singleton observations; the control group observations are included three times (one time for each treatment group). The last rows report p-values of tests on the pre-trend and treatment effect parameters. Compared to Table 2, we have added test statistics for whether effects at event time -1 and 0 are equal. Standard errors in parentheses are clustered on students:

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A3. Stacked regressions: Effects of separation on test scores by event time. The effect at event time 0 is split into effects at event time -1 (separation occurred within a year after the test date) and event time 0 (separation occurred within a year prior to the test date)

Event time (e)	(1) Control group: Treated at grade 10	(2) Control group: Treated at grade 10 or never
-6	0.0179	0.0288***
	(0.0178)	(0.0102)
-4	-0.00447	0.00343
	(0.00818)	(0.00575)
-2	0	0
	(.)	(.)
-1	-0.0174**	-0.0243***
	(0.00809)	(0.00617)
0	-0.0403***	-0.0411***
	(0.00802)	(0.00608)
2	-0.0533***	-0.0591***
	(0.0115)	(0.00691)
4	-0.0663***	-0.0627***
	(0.0195)	(0.0120)
N	160,435	2,785,063
Children	39,146	352,225
p-values of tests:		
All pre-trends 0	0.195	0.017
Equality for e ∈{-1,0}	0.013	0.038
Equality for e ≥ 0	0.258	0.039
Equality for e ≥ -1	0.002	0.000

Note. Results from stacked regressions of event study models controlling for individual fixed effects, grade fixed effects and calendar year of test fixed effects. The reference category for event time is -2. Compared to the models in Table 3 of the main analysis, we allow the effects at event time 0 to vary by whether separation occurred before or after the test. Thus, event time 0 in the main analysis represent observations where separation occurs from about 1 year prior to the test to about 1 year after the test. In the models of this appendix table, these observations are split up into event time -1 if separation occurs after the test, and event time 0 if separation occurs before the test (or at the same date). Otherwise, the specification is the same as in Table 3. The number of observations and children is excluding singleton observations; the control group observations are included three times (one time for each treatment group). The last two rows report p-values of tests on the pre-trend and treatment effect parameters. Standard errors in parentheses are clustered on students:

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A4. Robustness checks taking account of "anticipation effects": Effect of separation in grade 6 on test scores in grade 6 and 8 compared to test scores in grade 2

Event time (e)	(1) Control group: Treated at grade 10	(2) Control group: Never treated	(3) Control group: Never treated
0	-0.0303	-0.0350***	-0.0362***
	(0.0201)	(0.0114)	(0.0114)
2			-0.0632***
			(0.0157)
N	16,058	272,406	345,745
Children	8,029	136,203	137,571

Note. Results from linear regressions of models controlling for individual fixed effects, grade fixed effects and calendar year of test fixed effects. The table shows estimates of the effect of separation in grade 6. The reference category for event time is -4 (instead of -2 as in the main analysis) to take account of possible "anticipation effects", that is, the possibility of a prolonged period of escalation of social problems or conflict in the family leading up to parental separation. In model (1) the control group consists of children experiencing parental separation in grade 10. In models (2) and (3) the control group consists of the never treated. The number of observations is rather small because we only use test score observations for grades 2 and 6 (models (1) and (2)), or grades 2 and 6 or 8 (model (3)). The number of observations and children is excluding singleton observations. Standard errors in parentheses are clustered on students: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

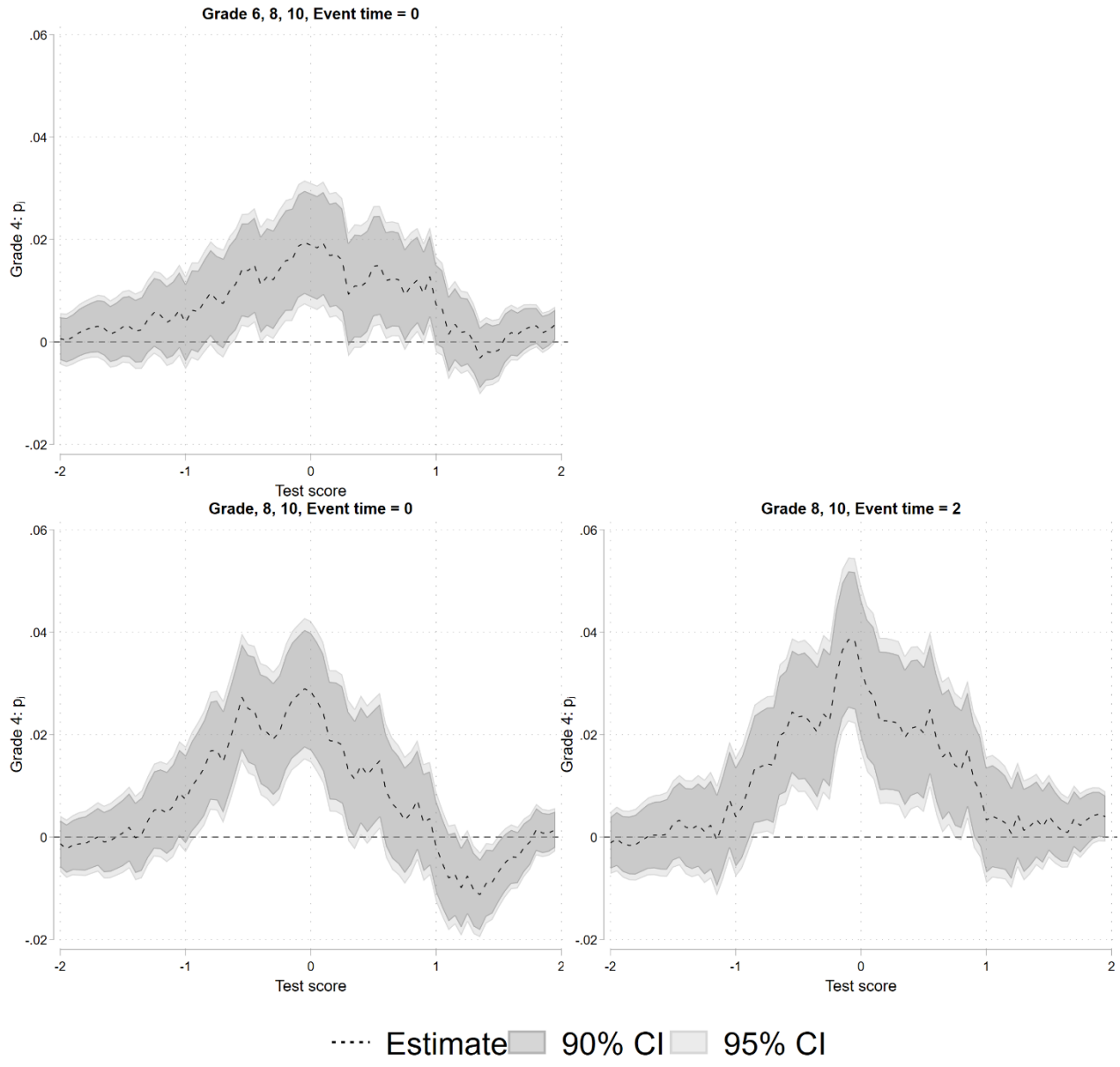
Table A5: Balancing tests of student and parent characteristics

	Estimate	Standard error
Male student	0.003	(0.004)
Young for grade	-0.003	(0.003)
Old for grade	0.000	(0.002)
Mother post-secondary degree	-0.001	(0.005)
Father post-secondary degree	-0.004	(0.004)
Maternal Income (€1,000)	-0.410*	(0.226)
Paternal Income (€1,000)	0.060	(0.235)

Note: The table shows estimates of the coefficient on the indicator for being above the threshold in RDD regressions for each variable (shown in the column head) on this indicator, the running variable, their interaction, and controls. All regressions control for grade of test and nearest test relative to parental union dissolution. Conventional RDD estimates from the rdrobust command in Stata. Triangular kernel, first order polynomial, and automatic bandwidth selection. Parental yearly total income (including public transfers) is deflated by the CPI (base 2015) and measured in €1,000 (exchange rate 7.44 DKK/€).

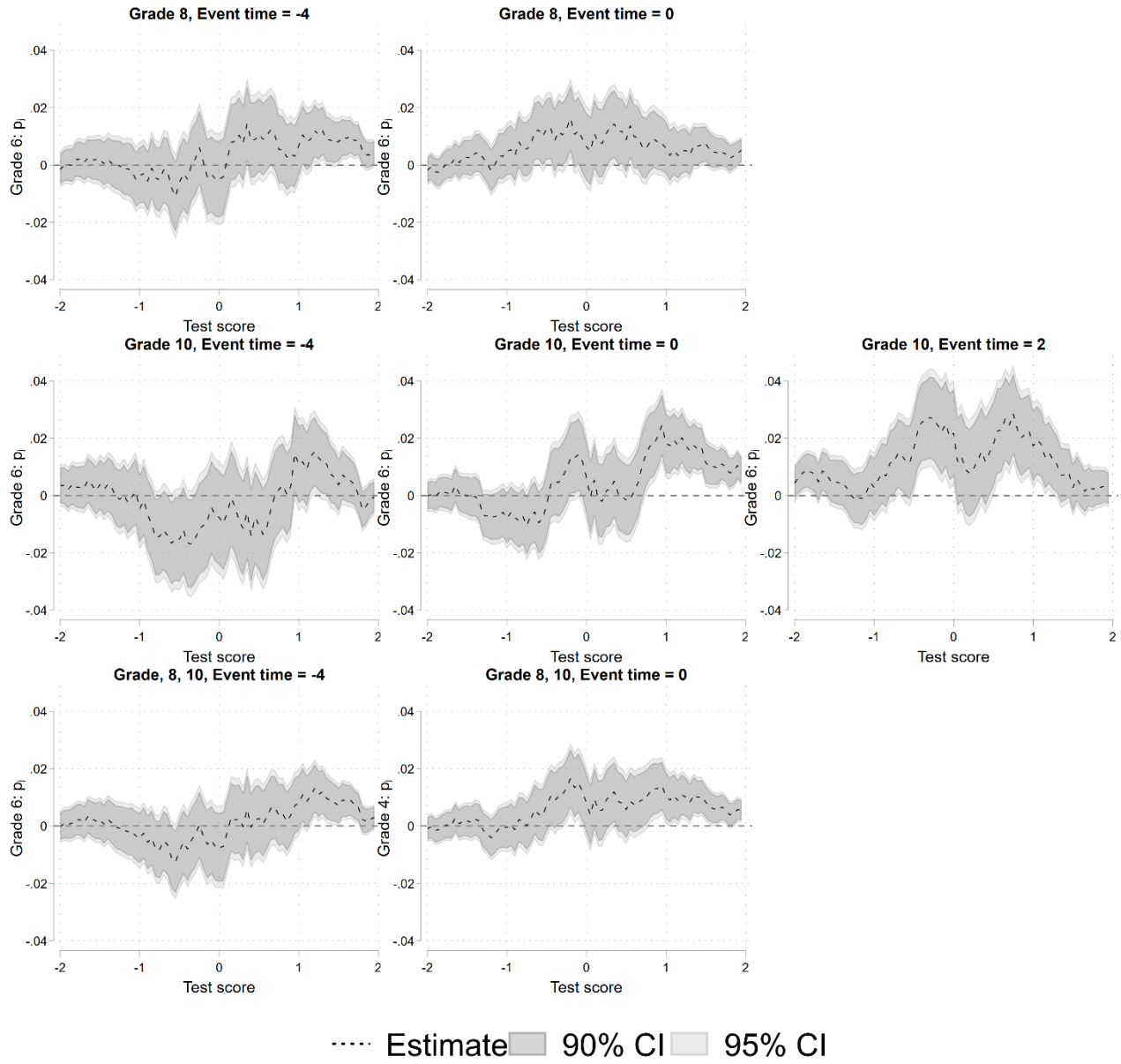
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure A1. Effect of union dissolution in grade 4 on the probability of scoring below a given level across the distribution of test scores (comparison group treatment timing and event time in subplot title)



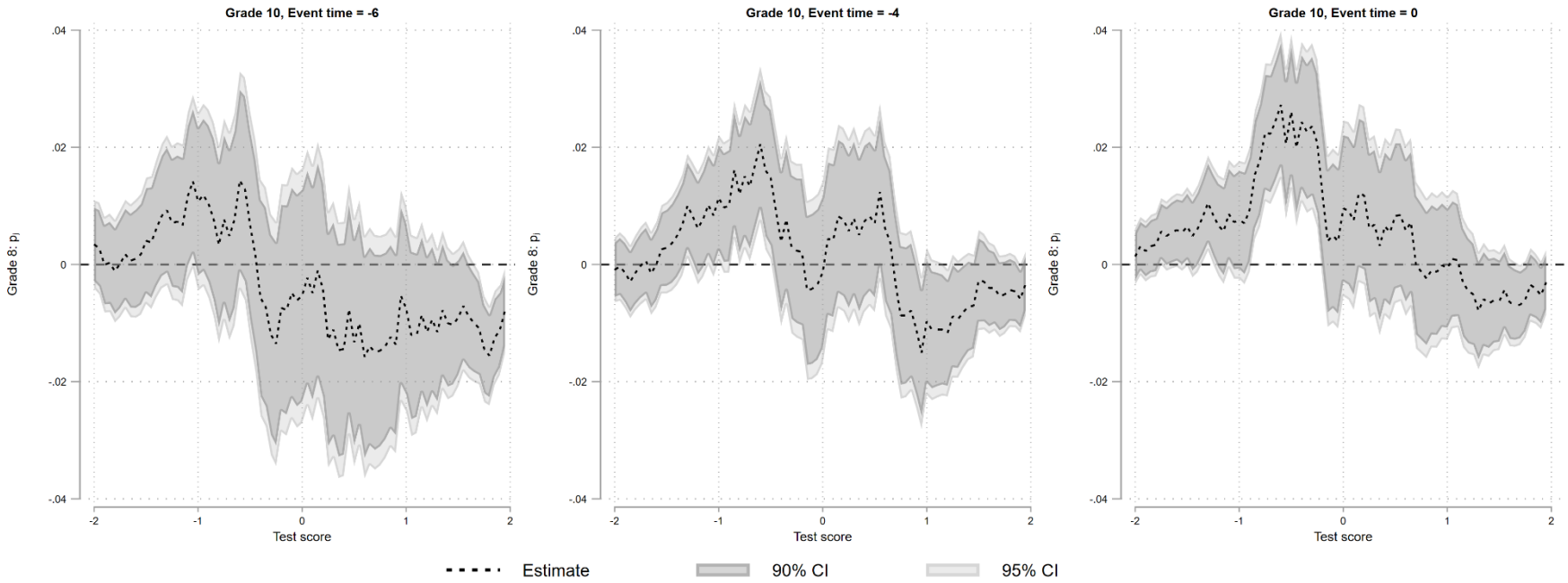
Note. The graphs show the effect of separation in grade 4 on the probability of scoring below the test score level indicated on the x-axis. Each subgraph shows effects for a given event time and a given comparison group consisting of later treated children. The upper subgraph shows effects at event time 0 using the largest possible control group of later treated (those experiencing separation at grades 6, 8 or 10). The two lower subgraphs show effects at event times 0 and 2 using the largest possible control group of later treated for event time 2 (those experiencing separation at grades 6 or 8). Confidence intervals based on robust parametric standard errors.

Figure A2. Effect of union dissolution in grade 6 on the probability of scoring below a given level across the distribution of test scores (comparison group treatment timing and event time in subplot title)



Note. The graphs show the effect of separation in grade 6 on the probability of scoring below the test score level indicated on the x-axis. Each subgraph shows effects for a given event time and a given comparison group consisting of later treated children. The upper subgraphs show effects at event time -4 and 0 using those treated at grade 8 as control group. The middle panel shows effects at event times -4, 0 and 2 using those treated at grade 10 as control group. The two lower subgraphs show effects at event times -4 and 0 using the largest possible control group of later treated for event time 0 (those experiencing separation at grades 8 or 10). Confidence intervals based on robust parametric standard errors.

Figure A3. Effect of union dissolution in grade 8 on the probability of scoring below a given level across the distribution of test scores (comparison group treatment timing and event time in subplot title)



Note. The graphs show the effect of separation in grade 8 on the probability of scoring below the test score level indicated on the x-axis. The three subgraphs show effects for event times -6, -4 and 0 using as comparison group children treated at grade 10. Confidence intervals based on robust parametric standard errors.