Consequences of Parental Divorce on Development of Cognitive Skills and Non-cognitive Traits in Childhood

Hyun Sik Kim

In this paper, we come up with three-stage estimation model to recover impacts of parental divorce on development of children's cognitive skills and non-cognitive traits. Using pre-, then-, and post-divorce framework, we can disentangle complex dimensions affecting children of divorce to more detailed extent. The Early Childhood Longitudinal Study-Kindergarten Cohort 1998-99, a multi-wave longitudinal data set provides an invaluable opportunity to assess the three-stage model. Abundant and more reliable measures on cognitive skills and non-cognitive traits as well as rich set of covariates open unchartered yet underutilized opportunity to appraise several hypotheses involved in most hotly debated areas in social sciences. To evaluate parameters of interest more rigorously, we employ stage-specific ordinary least square, counterfactual matching estimator, and piece-wise growth curve model. In general, we fail to detect statistically significant pre-divorce effect and total divorce effects as defined in this study across all developmental domains. Closer look at varied domains and stages reveals that impacts of divorce realize its influence not homogenously across the whole domains of but heterogeneously on specific and selective areas of children's outcomes in certain stages. Under some combination of developmental domains and stages, we find negative effects of parental divorce even after taking account for selection factors fulfilling their forces from or probably before dissolution process began.

1. Research Interest

Dominant portion of divorce literature has demonstrated adverse effects of parental divorce[1] on children's development (Amato & Keith 1991; Hetherington 2003; Cherlin et al. 1998; Wallerstein & Lewis 2004). In two authoritative meta-analyses, for example, Amato showed that cumulated evidence support the view that, compared to children with continuously married two-biological parents, those with divorced parents were disadvantaged in comprehensive domains of life chances: schooling outcomes, cognitive skills, psycho-social well-being, social relations (Amato & Keith 1991; Amato 2001). Moreover, recently published work has shown that these negative consequences did not diminish even in recent era characterized by more generous acceptance of divorce (Sigle-Rushton et al. 2005; Amato 2000).

Recent stream of research, however, has questioned the traditional null hypothesis of homogenous negative outcomes and its empirical evidence. Among many empirical and theoretical challenges, notable are selection perspective, observations on remarkable resilience of subpopulation, and more nuanced approaches on genuine effects of divorce per se due to possible confounding by other family processes preceding and following divorce such as marital discord and remarriage (Cherlin et al. 1991; Hetherington 1979; Amato & Hohmann-Marriott

[1] As is the case in the literature, we do not distinguish separation from divorce and consistently use more generic term divorce throughout this paper unless otherwise noted.

2007; Kelly & Emery 2003). For example, partly due to absence of appropriate data, whether preceding marital conflicts between parents are more responsible for children's outcomes or whether there are distinctive effects in dissolution process has not been explicitly addressed. Moreover, what characteristics of children of divorce and their related life world contribute to malleable resilience during and after divorce begs further investigation.

Up to date, rigorous study design incorporating these complex features in effects of parental divorce on involved children has been rarely conducted. Some authors recognizing and emphasizing divorce as a process failed to build statistical models appropriate for the conception (Cherlin et al. 1991; Morrison & Cherlin 1995). To bridge this gap in the literature, we evaluate three-stage effects of divorce on cognitive skills and non-cognitive traits in childhood: distinct as well as combined impacts during pre-, then-, and post-divorce period. To attain these goals, we use nationally representative prospective longitudinal data and advanced longitudinal statistical techniques. More specifically, we tap into the Early Childhood Longitudinal Study-Kindergarten Cohort 1998-99 (ECLS-K). Because ECLS-K traced children from kindergarten to 8th grade and measured rich set of family backgrounds variables, it provides an unmatched opportunity to disentangle several competing hypotheses involved in causal inferences of the divorce effects. Even better, ECLS-K contains cognitive skills measures calibrated to most up-to-date statistical techniques and diverse measures on behavioral problem scales assessed by teachers rather than parents who are more likely to report biased assessment on children's outcomes due to a priory concerns on their decision.

From methodological points of view, traditional ordinary least square (OLS) regression framework seems to be limited in causal inferences of the divorce effects. In particular, multivariate approach assuming balanced covariate set across the treatment and control group is quite burdensome for the current study because we apply such a strict definition of divorce that there are few observations on children of divorce. With a large number of the control group but relatively small number of the treatment group, matching estimator is recommendable statistical technique to recover parameters of primary interest (Rubin 1973; Smith 1997). In addition, our estimate of interest, the average treatment effect on the treated (ATT), provides attractive interpretation appropriate for our study purpose: ATT recovers average difference between realized developmental outcomes of children of divorce and counterfactual outcomes of those children had there parents remained married (Heckman et al. 1998; Heckman and Navarro-Lozano 2004). These OLS and matching estimator appear to be inadequate, however, in the face of our multi-wave longitudinal data. Concerns are with statistical inference rather than point estimates in stage-intersecting parameters due to observed and unobserved correlation within children across the time domain (Raudenbush & Bryk 2002; Singer & Willett 2003). To overcome these shortcomings, we supplement stage-specific and combined OLS and matching estimates with those from piece-wise growth curve model. Even though it is not widely used in sociological literature, piece-wise growth curve model turns out to be an invaluable tool for the current study. This is so especially because we are interested not only in distinct effects in several phases along the time dimension but also in their total effects aggregated over the whole time line.

2. Conceptual frameworks

Current paper aims at contributing to the literature by 1) taking into account selection argument as rigorously as possible, 2) evaluating phase-specific divorce effects, namely, the pre-divorce effect, the then-divorce effect, and the post-divorce effect, and 3) integrating those three-stage divorce effects into one unified total effects. We also cover comprehensive array of developmental domains: cognitive skills indexed by mathematics and reading test scores and teacher-assessed non-cognitive traits instrumented by interpersonal social skills, externalizing and internalizing behavior problems. To facilitate theoretical discussion and underlying motivations, we present the following figure which schematizes hypothetical developmental trajectories of children of divorce and those with continuously married biological parents. In the plot, x-axis refers to time frame and y-axis positive developmental outcome, say, math test scores. A set of the indices T s refers to the observation time points and D denotes the time of divorce which occurs between T_2 and T_3.

[Figure 1 about here]

In the graph above, developmental trajectory of children in intact family is supposed to move along the upper solid line of $P_{01} \rightarrow P_{02} \rightarrow P_{03} \rightarrow P_{04}$. In contrast, children of divorce may trace along lower solid line of $P_{11} \rightarrow P_{12} \rightarrow P_{13} \rightarrow P_{14}$. The latter trajectory reflects theoretical positions accounting for selection argument, measurement error in locating exact time of surfacing marital discord, negative pre-divorce effect, unfavorable then-divorce effect but identifiable resilience in post-divorce period.

There are three prominent reasons why children of divorce may trace the line from P_{11} to P_{12} instead of P_{01} to P_{02}: negative selection as well as positive selection, limitation in time measure on emerging marital strains, and adverse influences in pre-divorce period. Some selection variables such as children's psychological predisposition or frailty may have caused less favorable growth even before parents' marital conflict materialized (Amato 2001; Cherlin et al. 1991). In the opposite side, children with continuously married parents are more likely to enjoy supportive environments as far as parents-level selection mechanisms are operating. We would end up attributing unwarranted portion of selection effects to divorce effects, for instance, if parents destined to remain intact were originally more caring for and responsive toward their loved one so that those parents would be actively involved in and devote a large amount of effort to development of their offspring. Note that these selection factors are relevant only for the difference in the intercept, that is, the gap between P_{01} and P_{11} in so far as we condition on those factors in estimating the pre-divorce effect.

Or, even under the strong assumption of no such selection problems, we face measurement error regarding determination on the exact date when parental conflict emerged due to marital discord (Morrison & Cherlin 1995). If marital discord predated T_1, then measurement errors would put initial level of children's development onto P_{11}. In this case, by conditioning on measures from T_1, we may bias the pre-divorce effect to the positive direction. In other words, failure to adjust for negative outcomes that might have emerged from the initial stage of parental conflict would understate overall negative effects attributable to pre-divorce phase. However, if marital discord postdated T_1 conversely, then we would not incur such a bias. Given that our empirical data have one-year interval between T_1 and T_2, the former scenario would be more likely.

Divorce literature appears to agree to the point that even before formal divorce process is set in, relationship between couples destined to divorce is characterized by marital conflicts and, thus, their child(ren) is exposed to risk of developmental setback (Amato & Booth 1997; Cherlin 2008). Therefore, this perspective of negative pre-divorce effect is based on two premises: intense marital conflicts and their negative influences on involved child(ren)'s development. Even though recent research indicates reverse-causation hypothesizing child effects on marital conflicts, existing evidence seems to support the second argument (Jenkins et al. 2005; Emery 1982). For instance, in their classic book, Amato and Booth reported that children with two-parents maintaining conflict-ridden relationships were no less prosperous in psychological well-being and problem behaviors than those of divorced parents (Hetherington 1979; Amato & Booth 1997). The first argument is, however, not well established yet. Indeed, Amato showed that "sizable" number of divorce was "good enough marriages" in which marital discord was not readily noticeable (Amato 2002). Nevertheless, we can hypothesize negative pre-divorce effect, on average, as far as conflict-free marriages ending up with divorce do not provide positive effect enough to compromise negative force of conflict-prevalent marriages.

Figure 1 also illustrates control-away bias if we fail to consider pre-divorce process under the assumption of existence of noticeable pre-divorce conflicts and their negative consequences. Namely, if we estimated divorce effects using measurement from T_2 as baseline control variables, as mostly did, we would underestimate negative total divorce effects because already present marital conflicts must have decreased positive outcome in T_2, artificially diminishing unfavorable development tracing back to T_1. That practice of controlling T_2 outcomes, however, is valid as far as we are concerned with the then-divorce effect defined as difference in outcomes in the period spanning from T_2 to T_3.

Ample evidence has been cumulated supporting relatively deteriorating child outcomes in the divorce stage (Amato 1993; Lansford 2009). To name just a few theoretical mechanisms: continuing conflicts between divorcing parents (Hetherington 1979; Emery 1982), emotional troubles or lack of resources of divorced parents in adjusting to new environment leading to possible parenting problems (Cooper et al. 2009), economic hardship due to sudden drop of family income (Morrison & Cherlin 1995; Peterson 1996; McLanahan & Sandefur 1994); geographical relocation and school transfer following divorce (Astone & McLanahan 1994).

Notwithstanding such firm evidence, we can not exclude some possible routes through which parental divorce might contribute to child's well-being to the positive direction contrary to the conventional null hypothesis. Most notably, Amato and Hohmann-Marriott (2007) reported that the National Survey of Families and Households revealed two types of divorce, one of which was characterized by low-distress between the married couple and the other by high-distress. Interestingly enough, the former type reported loss of happiness after divorce while the latter scored increase in happiness. Those distinct, actually opposite, consequences for those involved adult couples suggest that children of divorce would be benefited from parents' decision to separate if parental marriages featured family dysfunctions and interparental conflicts before divorce. In addition, if there is no pronounced then-divorce effect or a child has already adjusted to parents' marital conflict to the extent that divorce per se does not pose any elevated negative risk, then developmental trajectory will move along $P_{12} \rightarrow P'_{13}$ instead of $P_{12} \rightarrow P_{13}$ [2].

[2] We note that the line $P_{12} \rightarrow P'_{13}$ is parallel to the line $P_{02} \rightarrow P_{03}$.

In the plot, we also include an assumption of measureable resilience after divorce was filed by painting the line $P_{13} \to P_{14}$ solid (Kelly & Emery 2003; Hetherington 2005). To avoid confusion on the usage of resilience in this paper, some discussion is necessary. To begin with, we acknowledge that resilience is a controversial concept and there is no unified framework to define it. Rather, there is ambiguity and uncertainty regarding 1) which characteristics should be considered responsible for resilience, for instance, a personal characteristics or family characteristics, 2) whether resilience should refers to bouncing back from adverse outcomes or relative absence of vulnerability, and 3) how to measure resilience, to name just a few (Kaplan 2005). In our context, do we believe resilience of children of divorce if there is no detectable pre-, then-, or post-divorce effect? Doing so seems to extend the concept of resilience to intractable extent because it misses the point that divorce effects should be examined but not assumed negative by definition. In this line of reasoning, we define resilience as bouncing back from previously negative outcomes as the route $P_{13} \to P_{14}$ shows.

Thus-defined resilience is not necessarily confined to the post-divorce period. For an extreme case, we can say resilience when there are noticeable negative outcomes in children of divorce in the pre-divorce period but, at the same time, we also witness significant positive turn-around in the then-divorce period. However, we do not observe that theoretical possibility in reality so that we restrict our term of resilience to the post-divorce period. In contrast, if developmental growth of divorced children was sustained apace with intact children without being aggravated, then former children would trace the line $P_{13} \to P'_{14}$ which is parallel with $P_{03} \to P'_{04}$. In this case, we do not say children are resilient. For an obvious reason, the antithetical scenario to resilience argument is that children might go along the line $P_{13} \to P''_{14}$ since there were unrestrained negative effects even after divorce (Wallerstein & Lewis 2004). If those detrimental mechanisms present in then-divorce period protracted into post-divorce period, considerably widening gap in developmental outcomes would be predictable.

In this regard, much attention has been devoted to family formation and its distinguishable effect on child's development after divorce because understandably family formation has been recognized as a source of confounding factor for identifying divorce effects per se. However, evidence is also mixed in relation to whether new family formation has overall negative effects. Using the Fragile Family and Child Wellbeing Study, for instance, Cooper and associates found that not only family transition type but also number of transitions mattered for maternal parental stress (Cooper et al. 2009). On the opposite side, however, Thompson and her colleagues found that mothering behaviors and mother-child relationship improved when mother was remarried or in partnership even though time elapsed after divorce might matter in this case (Thompson et al. 2001). Recent two articles analyzing the same data set (CNLSY) approximately reached to the similar conclusion that even though there existed noticeable differences in children's outcomes hinging on number of family structure transitions, those differences nearly disappeared after introducing rich set of selection factors (Aughinbaugh et al. 2005; Fomby & Cherlin 2007).

3. Statistical strategies

Usual OLS regression framework appears to be not suited for our current study because of multi-stage estimation strategies of the current study. To build up more cogent statistical models, let us refer to Y_t as a developmental outcome measured at time $t \in \{0,1,2,3,4\}$. We follow conventional practice in notation by writing capital letter for a variable and small letter for realized value of relevant variables. Likewise, $\mathbf{X}_t$ carries a vector of confounding variables observed at time t. D denotes the treatment variable evaluated unity if parents divorced and zero otherwise. Note that there is no subscript to the treatment variable because it is not time varying variable. To assist understanding on the statistical models, we present a causal directed acyclic graph embodying underlying causal relationships of the current study in the following figure (Pearl 2000).

[Figure 2 about here]

It is quite straightforward to model divorce effects in the period stretching from spring first grade to spring third grade when parental divorce actually realized. To remove confounding bias, it is necessary to condition on the covariate set $\mathbf{X}_2$ to estimate the then-divorce effect. For more conservative estimates, we include in $\mathbf{X}$ all the cognitive skill and non-cognitive trait variables with one survey-time lagged. Also note that Y_2 acts as a confounding variable, acknowledgement of a child effect. In other words, failure of a child to live up to parents' expectation may influence parents' decision on marital dissolution (Schermerhorn et al. 2007; Cui et al. 2007). To materialize these theoretical points, a formal model may be specified as:

$$Y_3 = \beta_{21} D + \beta_{22} Y_2 + \mathbf{X}_2^T \boldsymbol{\alpha}_2 + \varepsilon_3 \qquad 1$$

where ε_3 stands for error term. As usual, the superscript T to $\mathbf{X}_2$ means vector transpose because a vector is written as a column vector. Consequently, $\boldsymbol{\alpha}_2$ is a parameter vector associated with $\mathbf{X}_2$ in which the leading row consist of unity to accommodate the intercept term. Under Equation 1, our interest center around β_{21}, the parameter for the then-divorce effect.

How can one estimate the pre-divorce effect? It should be immediately clear that it is literally impossible to estimate the pre-divorce effect upon recognizing that by the pre-divorce effect, we want to predict Y_2 using D even though the latter was in actuality generated by the former, a sheer contradiction. This statistically nonsensical enterprise, however, has a theoretically robust ground. As discussed in the theoretical part, most divorce exhibited marital strains and interpersonal conflict affecting involved child(ren) and, therefore, omitting this pre-divorce process would bias total divorce effects to the favorable direction. However, it is also crucial to adjust for other covariates confounding this pre-divorce effect. Under this reasoning, we will estimate the pre-divorce effects by

$$Y_2 = \beta_{11} D + \beta_{12} Y_1 + \mathbf{X}_1^T \boldsymbol{\alpha}_1 + \varepsilon_2 . \qquad 2$$

We caution that the estimate β_{11} do not have any causal meaning unlike estimates from other stages. In that sense, the estimate provides only descriptive conjecture on the pre-divorce effect at best.

Lastly, how one can plausibly recover the post-divorce effect? It may be tempting to use the following equation and claim β'_{31} to be a relevant estimate on the post-divorce effect.

$$Y_4 = \beta'_{31} D + \beta'_{32} Y_3 + \mathbf{X}_3^T \boldsymbol{\alpha}'_3 + \varepsilon'_4 . \qquad 3$$

However, this approach controls away possible contributions of divorce to the post-divorce effect. Figure 2 above shows that, for instance, there is a path from D to Y_4 via Y_3, which means that controlling Y_3 block the path, eventually committing a bias. Conditioning on Y_3 and $\mathbf{X}_3$ as with Equation 3, we will estimate the sole direct path from D to Y_4 rather than overall effects. Instead, more robust estimate can be obtained by estimating

$$Y_4 = \beta_{31} D + \beta_{32} Y_2 + \mathbf{X}_2^T \boldsymbol{\alpha}_3 + \varepsilon_4 \qquad 4$$

in which β_{31} realizes the combined effects of then- and post-divorce effects. Therefore, to get the post-divorce effect, we subtract the then-divorce effects from the combined effects, say, $\beta_{31} - \beta_{21}$ derived from Equation 4 and Equation 1 consecutively. On the other hand, total divorce effects can be summarized by adding the pre-divorce effect to the combined effects, namely, $\beta_{31} + \beta_{11}$. Nevertheless, we argue that Equation 3 also provides useful estimates particularly in policy context. Provided that predicting parental divorce is a speculative job at best, we can only observe children of divorce after they went through it. Under this real setting, policymakers are interested in how much the divorce effects last conditional on the currently observed covariates. Equation 3 exactly supplies an answer to that question so that we also report estimates based on Equation 3.

One of the problems applying OLS estimator separately to the estimands specified through Equation 1, 2, 3, and 4 is that even though we can obtain point estimates and p-values for stage-specific parameters, statistical inference is not available for those estimates across stages such as total divorce effects. Most straightforward way to overcome this shortcoming is to bootstrap subsamples with replacement, apply OLS to each subsample for a specific stage, obtain all the estimates of interest for a subsample, calculate standard errors using all the estimates and test the null hypothesis in the conventional level of α equal to .05 (Efron & Tibshirani 1993; Lohr 1999). We randomly subsample 90 times our analytical data generated by list-wise deletion to maintain consistency with weighting methods employed to take account of longitudinal attrition as we will discuss shortly. However, we also mention that the bootstrap method would give a little imprecise standard errors particularly when it comes to aggregate estimates due to correlated nature of our observations along the time line (Raudenbush & Bryk 2002; Singer & Willett 2003).

In this paper, we also explore counterfactual method via a matching estimator because of its easiness in estimation and successful application in previous research. Using the propensity score as a weight, for example, Amato (2003) found that offspring of divorced parents are more likely to have lower self-assessed psychological well-being, more prone to marital discord, and less likely to maintain a good relationship with father. From estimation point of view, matching method may provide quite robust estimates without extrapolating treatment effects beyond common support of confounding variables (Rubin 1973; Smith 1997). Especially because there are small number of observations on the treatment group (N=142) and substantially larger number of the control group (N=3,447) in our data set, we concentration on estimation of the average treatment effect on the treated (Heckman et al. 1998; Heckman and Navarro-Lozano 2004). Another advantage in the estimation of ATT lies in its attractive interpretation. As its mathematical form suggests, the effect refers to the average difference between realized developmental outcome of a divorced child and counterfactual outcome the same child if the child did not experience parental divorce.

To develop formal discussion, we start from the basic fact that one student can never live with the family both divorced and continuously married at the same time. Thus, plausible estimate on divorce effects is averaged estimate in the level of population only after advertent and inadvertent selection has been made. For a more formal discussion, let $Y = Y(0)$ if a student lives with continuously married parent while $Y = Y(1)$ for a child of divorce. In addition, let $D = 0$ if parents "choose" to remain married while $D = 1$ for divorce. Then, more convenient notation would be $Y = DY(1) + (1 - D)Y(0)$. If our interests are restricted to estimating the treatment effect on the treated, $E[Y(1) - Y(0) \mid D = 1]$, bias from estimating the effect using $E[Y(1) \mid \mathbf{X}, D = 1]$ and $E[Y(0) \mid \mathbf{X}, D = 0]$ would be

$$\begin{aligned}
&[E[Y \mid \mathbf{X}, D = 1] - E[Y \mid \mathbf{X}, D = 0]] - E[Y(1) - Y(0) \mid \mathbf{X}, D = 1] \\
&= [E[Y(1) \mid \mathbf{X}, D = 1] - E[Y(0) \mid \mathbf{X}, D = 0]] - [E[Y(1) \mid \mathbf{X}, D = 1] - E[Y(0) \mid \mathbf{X}, D = 1]] \\
&= E[Y(0) \mid \mathbf{X}, D = 1] - E[Y(0) \mid \mathbf{X}, D = 0] \qquad 5
\end{aligned}$$

where $\mathbf{X}$ is a set of confounding variables as before.

Under the assumption of the unconfounded treatment assignment conditional on a set of the observed confounding variables $\mathbf{X}$, namely,

$$D \amalg Y(d) \mid \mathbf{X} : d \in \{0,1\}, \qquad 6$$

$E(Y(0) \mid \mathbf{X}, D = 1) = E(Y(0) \mid \mathbf{X}, D = 0)$ so that bias will disappear. Or, one might prefer weaker assumption that $E(Y(0) \mid \mathbf{X}, D) = E(Y(0) \mid \mathbf{X})$ for the bias removal of the average treatment effect on the treated (Heckman et al. 1998; Hirano et al. 2003). Under this setting, the average treatment effect on the treated can be obtained by the iterated expectation formula:

$$E[Y(1) - Y(0) \mid D = 1] = E_{\mathbf{X}}[E[Y(1) - Y(0) \mid \mathbf{X}, D = 1]]. \qquad 7$$

In their classical article, Rosenbaum and Rubin (1983) argued that if the condition of Equation 6 is satisfied and there is no confounding variable perfectly classifying the treatment variable such that the conditional probability of getting the treatment given the set of observed confounding variables, so-called, the propensity score, $p(\mathbf{X})$, stays away from zero and unity, conditioning only on the propensity score is sufficient to remove potential bias in estimating the average treatment effect (Rosenbaum & Rubin 1983; Heckman & Navarro-Lozano 2004; Hirano et al. 2003). It can be shown, however, that currently available matching estimator based on the propensity score does not necessarily give unbiased estimates if matching is not exact all across covariates' values (Abadie & Imbens 2002; Abadie et al. 2006). Due to this potential problem in the propensity score matching method, considerable scholarly efforts have been devoted toward attainment of reasonable balancing in values of covariates among matched observations (Dehejia & Wahba 2002).

Due to this consideration, we use "GenMatch" routine in R which was developed by Sekhon (Sekhon forthcoming; Diamond & Sekhon 2006). In a nutshell, the genetic matching algorithm is to stochastically find a weighting matrix to minimize Mahalanobis distance among confounding variables so that it can achieve tighter balancing. In a mathematical notation, the genetic matching search an weight matrix $\mathbf{W}$ to minimize the generalized Mahalanobis distance

$$d(\mathbf{X}_i, \mathbf{X}_j) = [(\mathbf{X}_i - \mathbf{X}_j)^T (\mathbf{S}^{-1/2})^T \mathbf{W} (\mathbf{S}^{-1/2})(\mathbf{X}_i - \mathbf{X}_j)]^{1/2} \quad 8$$

where $\mathbf{S}^{1/2}$ is the Cholesky decomposition of the variance-covariance matrix of $\mathbf{X}$. The GenMatch function also provides somewhat general matching methods such as caliper matching and one-to-many matching. We tried matching up to 9 control observations for a child of divorce but as expected, balancing statistics have been severely deteriorated beyond one-to-one match so that we report results from one-to-one match. For more interested readers, we invite to see Appendix B reporting balancing qualities for covariate values by number of controlled children matched to a treated child. We also report consistent estimates for the large sample variance developed by Abadie and Imbens in a recent series of papers, which is a built-in feature of GenMatch (Abadie & Imbens 2002; Abadie & Imbens 2006; Sekhon forthcoming).

To estimate the then-divorce effect, we condition observed values of confounding variables at the spring first grade to match the treatment group to the control group and compute the average treatment effect on the treated. Due to balancing requirement for matching estimators, matched sample for the then-divorce effect should not show difference in values of confounding variables between the treated group and the control group, which is the simple reason why we can not use the matched sample to obtain the pre-divorce effect. Thus, we condition the confounding variables collected at the spring kindergarten survey and capture pre-divorce effect using thus generated matched sample to get the pre-divorce effect. This procedure again highlights why our estimate on the pre-divorce effect do not possess causal interpretation. As to estimation of post-divorce effect, we also follow the reasoning outlined before. To repeat, concerns with controlling away potential effects partly generated by divorce refrain us from using covariates measured at the spring third grade so that we reuse those matched observations created for the then-divorce effect in order to come by combined effects of then- and post-divorce effect. Similarly, we produce post-divorce effect by subtract then-divorce effect from the combined effect. On top of that, we also present matching estimates similarly to Equation 3.

Our final model building strategy is based on piecewise growth curve model. As previous Figure 1 suggests, the current research problem can be viewed as estimating two different growth trajectories depending on parental decision to divorce. To formulate this perspective in mathematical language, let Y_{it} denote an developmental outcome of a child i at time t. Piecewise growth curve model is specified by

$$\begin{aligned} &\text{Level 1: } Y_{it} = \pi_{0i} + \pi_{1i}a_1 + \pi_{2i}a_2 + \pi_{3i}a_3 + \varepsilon_{it} \text{ for } t = \{1,2,3,4\} \\ &\text{Level 2: } \pi_{li} = \delta_{l0} + \delta_{l1}D_i + \delta_{l2}Y_{il} + \mathbf{X}_{il}^T \boldsymbol{\alpha}_l + u_{li} \text{ for } l = \{0,1,2,3\} \end{aligned} \quad 9$$

where $\varepsilon_{it} \sim N(0, \sigma_\varepsilon^2)$ and $u_{li} \sim MN(\mathbf{0}, \mathbf{\Sigma})$ in which N and MN denote normal and multivariate normal distribution respectively and $\mathbf{\Sigma}$ is a 4-by-4 covariance matrix allowing non-zero covariance in upper and lower diagonal cells. To represent piecewise growth curve correctly, a_1, a_2, and a_3 should be coded as the following table.

[Table 1 about here]

In Equation 9, parameters of critical interest are δ_{11}, δ_{21}, and δ_{31} that represent growth parameters in each period. This model also enables us to estimate total divorce effects and test the effects statistically by summing growth parameters across time dimension, say, $\sum_{l=1}^{3}\delta_{l1}$. Note that we do not include δ_{01} which stands for initial level of development primarily because it confounds selection effects and measurement errors in locating emergence of marital strains with divorce effects as we detailed in previous discussion. We use "PROC MIXED" routine in SAS for model estimation (Singer & Willet 2003). We encounter convergence problems in several models even after we try comprehensive set of initial values such as converged estimates from "lmer" function in R (Bates 2008). With the hope of normalization, we also try log transformation on the non-cognitive trait variables on the original metric of which minimum value was unity. In that case, we transform all the non-cognitive trait variables whether they are used as covariates or response variables. Though not all models converge in that case either, we see the transformation strategy work to some extent. We report statistical results from both natural metric and transformed metric.

Needless to say, how to adjust for longitudinal attrition should be a major concern for longitudinal data analysts. To ameliorate attrition bias, we adopt design-based method: weighting by longitudinal weight ("c1_6fp0") furnished by the data collector (Lohr 1999; Tourangeau et al. 2006). To obtain statistical inference for estimates on weighted data, Tourangeau et al. recommend the paired jackknife method using the replicate weights which are also equipped in the longitudinal data set. Since the recommended method means deriving standard errors from repeated estimates using the replicate weights as specified in standard textbooks, we follow the recommendation. For more complete report, we show both classes of estimates, one using unweighted data and the other using weighted data. To our knowledge, there is no consensus on how to use weights in matching estimators: whether weights should be used when Mahalanobis distance is calculated, how weights should be used when it comes to matching, and whether weights from the control group should be used in outcome estimation. We obtain optimal matched pairs without using longitudinal weights and repeatedly apply the replicate weights of the treatment group to both matched pairs so that a given pair shares the same longitudinal weight of the treatment group.

4. Data and measurement

1) Data

To implement preceding conceptual and statistical models, we utilize ECLS-K longitudinal data set. The ECLS-K is a nationally representative study with a multistage probability sample from the population of the 1998-99 kindergarten cohort (Tourangeau et al. 2006). With geographical areas being the primary sampling units, the NCES chose schools as the second-stage units from which students were sampled. The study consisted of the initial survey in the fall of kindergarten and six follow-ups [the spring of kindergarten (1998-99), the fall and spring of first grade (1999-2000), the spring of third grade (2002), the spring of fifth grade (2004), and the spring of eighth grade (2007)].

Among those available waves, we do not include the fall survey conducted in the first grade because only 30% subsample of eligible children was interviewed at the wave. Thus, the only children who have data before and after divorce are those whose parents divorced between the spring of first grade and spring of the third grade[3]. We choose the interview fielded in the spring of kindergarten as $T = 1$ survey. Two considerations are involved in this decision: test scores in fall semester may be dominated by summer schooling parameter and teacher-assessed non-cognitive traits are more prone to incur measurement errors due to short time observation on the involved children. As mentioned before, we also need covariates taken from the fall of kindergarten to construct reasonable piece-wise growth curve model. All in all, we concentrate our attention on five surveys: the fall and spring of kindergarten, the spring of first grade, third grade, and fifth grade, with the first one acting as a baseline. These survey rounds will be denoted by $T_t : t \in \{0,1,2,3,4\}$, respectively.

2) Measures

A. Treatment variable: divorce

We concentrate our attention on comparison between children who experienced parental divorce in the period from the spring of first grade to the spring of third grade and those whose parents stayed married throughout that period. Note that we include in the definition of children of divorce those children who also exposed to other family processes such as cohabitation or remarriage after divorce by not considering marital status after the third grade interview[4]. Also note that only those children who remained intact from the initial survey until the spring of third grade were eligible for control groups. More important, parents involved are biological two parents so that children with adopting parents or remarried parents are all excluded from analytical sample. This decision comes from the firm position for more rigorous evaluation on divorce effects per se. This rigorous definition of children of divorce may equalize support of covariates between two comparison groups and hopefully ameliorate selection bias in exchange for reduced sample size.

We used four variables to operationalize divorce: marital status and parents' types in household at T_2 and T_3. Only those students whose biological two parents were married and lived together at T_2 were eligible for the analytical sample. Among those analytical sample children, control group is made up of children who enjoyed the same two conditions at T_3. In contrast, we define the treated children to be those whose parents were not two biological parents at T_3 regardless of marital status in order to include separation cases into divorced children. We, however, exclude widowed children at T_3 from the definition of divorced children to avoid misclassification problem (Emery 1982).

---

[3] As of the analysis for this paper, the eighth grade data are not available.

[4] One may argue that to avoid confounding with other family process after divorce, we need to compare children whose divorced resident father or mother stayed unmarried or unpartnered with those whose two biological parents stayed married until the fifth grade interview. It turns out, however, that this approach brings about more complications rather than solutions by introducing endogenous selection bias, a.k.a., Berkson's paradox, or explaining-away bias (Elwert & Winship 2008; Pearl 2000). Recalling our DAG displayed in Figure 2, let's suppose that we are interested in predicting an outcome measured at T_3 using covariates collected at T_2. Notice that marital status (single status for divorced parents as well as married status for non-divorced parents) at T_4, thus, probability to include in the sample is a variable most likely correlated with divorce and quite plausibly with an outcome variable observed at T_3 as well. Then, controlling marital status, in other words, selecting observations stayed married or single at T_4 give rise to artificial correlation between divorce variable and the outcome variable leading to overestimation or underestimation of parameters of interest depending on the nature of relationship between the divorce and outcome variables. For those who are concerned with confounding with subsequent family building process, we hope that two-year duration of post-divorce period would not pose serious risk.

B. Outcome variables

Regarding the measurement of variables, how to index the "development" of cognitive skills remains a critical issue. In particular, there is a problem of test score metrics: among five metrics provided by the ECLS-K public data, it is recommended to use proficiency probability scores for longitudinal cognitive development analyses (Tourangeau et al. 2006). However, proficiency probability scores consist of 9 dimensions for each subject. Due to its difficult implementation deriving from a multitude of dimensionalities, we use the Item Response Test (IRT) scale scores for this study. The IRT scale score can be interpreted as probabilistic scores with respect to the number of correct answers a student would have made if she were given all 153 questions in mathematics and 186 questions in reading.

For non-cognitive trait measures, we use three teacher-assessed social rating scales which are supposed to capture children's socio-emotional development. Originally there are five non-cognitive measures encompassing 1) approach to learning, 2) self-control, 3) interpersonal social skills, 4) externalizing problem behaviors, and 5) internalizing problem behaviors. Each sub-measure consists of six, four, five, five, and four items respectively. Scale of each item range from "never (1)" to "very often (4)". Split-half reliability for sub-measure reveals quite good reliability for all measures usually hovering over 0.8 (Tourangeau et al. 2006). Among those five measures, however, we use only the last three variables. One reason to do so is that first two variables are highly correlated with test score measures so that no much new information can be gained from modeling those two variables. Furthermore, due to large correlation, estimates from multivariate approach may become unstable especially when we condition on test score measures.

In addition, we also note that two more individual items were added from the spring of third grade in addition to original 24 items, one for externalizing problem behaviors and the other for self-control. Because we use mean level of each sub-measure[5], added items are not likely to pose any serious problem to the analyses presented here. All scales are relocated to have range from 0 to 3 with high values denoting high frequency of a relevant behavior. For example, high score in interpersonal social skills indicates a student's good skills in interpersonal exchanges while high score in externalizing problem behaviors suggests high frequency of student's problematic behaviors.

C. Confounding variables

Selection Factors To control parent-level selection, we include a measure on whether parents were married when a student were born. Although this is an imperfect measure given the possibility that the focal child may be not the only child, our hope is to capture marital selection problems through conditioning marital status at the birth of the focal child (Carlson et al. 2004). We also use a measure on mother' psychological well-being evaluated at the baseline survey. ECLS-K collected 12 self-assessed items on psychological well-being. For example, one item asked: How often during the past week, do you feel depressed? Each item has values from 1 (never) to 4 (most of the time). Alpha reliability of those 12 items amounts to 0.857 when total available samples are used. We calculate average of those 12 items to subtract one in order to have range of zero to three and use it as a control variable with continuous scale. Further, we also include self-assessed global happiness on marital relationship. In the spring semester of kindergarten, the questionnaire asked how parents responded to the question about the

[5] The data collector releases only summary statistics of average for entire classes of sub-measures.

relationship with spouse. Response space consisted of not too happy (=1), fairly happy (=2), and very happy (=3). Since it is well known that this variable is highly skewed to the left, we include it as categorical variable.

Other confounding factors On top of those variables, we also consider basic demographic variables: age in month as of June, 2000 (Hetherington 1979; Emery 1982), gender (Cherlin et al. 1991; Morrison & Cherlin 1995), race/ethnicity (Bulanda & Brown 2007), number of siblings, urbanity (Gautier et al. 2009), and geographical region (Glenn & Shelton 1985) as well as school move between two adjacent waves (Astone & McLanahan 1994; Boyle et al. 2008). Needless to say, it is critical to include socio-economic status variable (Amato & Booth 1991; Cherlin 2008; McLanahan & Sandefur 1994). Among various measures to index socio-economic status, we just include socio-economic status index provided by data collector which was calibrated to average of five family background variables (father or father figure's education and job prestige, mother or mother figure's education and job prestige, and household income), each of which were normalized to have mean zero and unit standard deviation before they were summed (Tourangeau et al. 2006).

5. Descriptive Statistics

Table 2 in the following page shows descriptive statistics for treatment, outcome and lagged outcome variables both without and with weights in the analytical sample. Nonetheless burdensome amount of information, we provide descriptive statistics separately for the two treatment groups because one of our estimator is matching method picking up the average treatment effect on the treated such that understanding on the treatment group are essential for evaluation of statistical estimates. We also present descriptive statistics for other confounding variables in Appendix Table A. Among total 3,589 children, only about 4% experienced parental divorce between spring semester of first grade and spring semester of third grade. When weighted, the estimate slightly increases to more than 6%, which suggests that more children of divorce might have been lost to follow-up until the spring fifth grade compared to children of continuously married parents.

[Table 2 about here]

Regarding outcome variables such as mathematics and reading test scores indexing cognitive skills, we find that 1) average difference were already present in the initial survey between two treatment groups, 2) the difference seems to have increased as time went on, but 3) whether the elevated difference would be translated into compelling evidence on diverging difference is questionable because population standard deviation also tended to increase noticeably with time. To illustrate, average differences in math test score at the fall of kindergarten was 3.8 and 3.4 but they went up to 7.2 and 10.4 by the spring of fifth grade in unweighted and weighted sample respectively. However, population standard deviation also more than doubled in the period. These findings are consistent in both weighted and unweighted sample except that mean estimates in weighted sample consistently show moderately shrunken values in contrast to those in unweighted sample, which is most likely another indication of positive and negative selection in longitudinal follow-ups.

As to non-cognitive traits variables, similar observations hold true besides that there was no patterned increase or decrease in average levels of non-cognitive traits for specific population. When it comes to internalizing problems variable, however, we see a readily noticeable increase in difference between two treatment groups in the period from T_2 to T_3 during which parental divorce had been filed. It is interesting to point out differences in population estimates of standard deviations between cognitive skill measures and non-cognitive trait measures by the treatment status. Patterned finding is that standard deviations coupled with minimum and maximum values in children with intact family tend to show wider distribution for cognitive skill measures but narrower distribution for non-cognitive trait variables. Put together, these findings suggest that there may be a strong selection effect operating from or probably before the baseline survey and impacts of divorce may realize its influence not on the whole domains of developmental outcomes but on more specific and selective areas of children's outcomes.

Appendix Table A suggests that there is no much difference in marginal distribution of basic demographic variables such as age and gender. Black children are somewhat over-representative in divorced population (Bulanda & Brown 2007). Interesting observations include dramatic change in distribution of urban location contingent on whether weight is given or not. When weighted, substantially more children of divorce resided in city but there is no distinguishable pattern otherwise. Given that previous report documenting higher divorce rates in urban areas, weighted estimates deem more reliable (Gautier et al. 2009). We also notice enhanced risk of school move in the population of divorced children especially around the time of divorce, which may be explained by the observed association between risk of geographical mobility and risk of divorce regardless of causal precedence (Glenn & Shelton 1985; Boyle et al. 2008).

Turning to selection related variables, somewhat noticeable difference emerges. For example, substantially higher percentage (13.4%) in the divorced family was not married at the time of the focal children's birth compared to that from the intact family (6.0%) though the difference gets smaller when the observations are weighted to 8.9% and 6.8% respectively. Also, parents were more likely to report marital dissatisfaction at the baseline survey if they would end up with marital dissolution. In addition, parents who would divorce in a few years reported a little higher score on psychological symptoms. However, no other variables show more dramatic difference between two groups than socio-economic variable. Especially, difference in the baseline is remarkable with intact family positioned in advantaged social strata. Nevertheless, it is barely surprising given repeated findings consistent with this observation (Cherlin 2008; McLanahan & Sandefur 1994).

To investigate these preliminary observations to more sophisticated extent, we now turn to more formal statistical models.

6. Statistical results

Table 3 through 7 displays results for the statistical models with cognitive skill variables and non-cognitive traits variables being response ones subsequently. For each response variable, we estimate three classes of statistical models: OLS, matching, and piece-wise growth curve model. For each statistical model, we fit weighted and unweighted sample. In unweighted sample estimation, we provide standard errors and their p-values based on asymptotic theory as well as those obtained by Bootstrap method and the paired Jackknife method. As mentioned already, the second class of estimates is useful for statistical inferences in aggregate estimates across stages. We retrieve standard errors for weighted samples with the paired Jackknife method following data collector's recommendation (Tourangeau et al. 2006; Westat 2007). Model names in the tables are designed to deliver which phase is concerned in a specific column. For example, $T_0 \rightarrow T_1$ means that the column contains estimates for outcome variables measured at T_1 with explanatory variables collected at T_0. Some comments on the estimates about $(T_2 \rightarrow T_4) - (T_2 \rightarrow T_3)$ are worth repeating as suggested in the discussion on statistical strategies. Particular downside of the effect flow in $T_3 \rightarrow T_4$ is that control variables measured at T_3 are likely to intercept realized divorce effects in previous stages so that genuine divorce effects may be underestimated or overestimated relying on relationships between divorce and covariates at T_3 together with those between covariates and outcome variables realized at T_4. To lessen this statistical problem, we assess the path $T_2 \rightarrow T_4$ and also provide the estimate subtracting the quantity on the line $T_2 \rightarrow T_3$ from the path $T_2 \rightarrow T_4$ for an alternative measure on the post-divorce effect. In the same vein of reasoning, we can construct total divorce effects with $T_1 \rightarrow T_2 \rightarrow T_3 \rightarrow T_4$ as well as $T_1 \rightarrow T_2 \rightarrow T_4$. However, note that piece-wise growth curve model can only put forth the causal path $T_1 \rightarrow T_2 \rightarrow T_3 \rightarrow T_4$.

[Table 3 about here]

For ease of presentation, we proceed from results on models predicting mathematics test scores [Table 3]. In general we find that point estimates suggest disadvantaged performance of divorced children as opposed to children with continuously married two biological parents. However, it is more or less surprising to find consistently positive estimates on the pre-divorce effect even though those coefficients fail to attain statistical significant except the weighted OLS model. One may speculate that this phenomenon may be related to pre-divorce version of "grow-up a little faster", a keen observation by Weiss who noted that children in single parent family often more matured compared to those with two-parents because single parent would share family works and responsibilities with children (Weiss 1979; Koerner 2006). However, primarily because it do not replicate in other developmental dimensions under current investigation, we do not further discuss it. Other than that, all the coefficients seem to be in accord with our theoretical prediction across two following phases.

While statistical test informs more conservative interpretation for the stage-specific effect, combined effects of the then-divorce effect and the post-divorce effect are statistically significant mostly within the conventional p-value of $\alpha = .05$, especially when the combined effects are defined by $T_2 \rightarrow T_4$. When we conceive it by $T_2 \rightarrow T_3 \rightarrow T_4$, statistical inference do not agree across weighting methods. Put this in perspective, simple random sample assumption across longitudinal waves gives statistically insignificant estimates while adjusting for attrition indicates statistically significant difference. Regarding total divorce effects, we see substantively different statistical inference depending on preferences in involved assumptions of each estimator: multivariate frameworks indicate statistical insignificance of negative influence by marital dissolution of parents while matching estimates suggest large and statistically significant effects. For instance, children of divorce were left behind, on average, 8.7 points in mathematics test scores compared to the counterfactual scores they would attain had their parents stayed married when we perceive total divorce effects in the framework of the path $T_1 \rightarrow T_2 \rightarrow T_4$ with weight considered. Our descriptive statistics recall that the quantity amounts to approximately a half of a standard deviation at T_4.

[Table 4 about here]

Estimates on reading test scores [Table 4] are characterized by closer correspondence to our expectation of negative divorce effects in point estimates. Yet imprecise point estimates bring about comprehensive failure in rejecting the null hypothesis of the zero effect. Unweighted OLS estimates suggest existence of a negative then-divorce effect and combined then- and post-divorce effect as well as total divorce effects but these results are too method-sensitive to be accepted wholeheartedly as a rigorous evaluation on one of the most hotly debated subject (Amato 2003). Likewise, substantially larger negative values in ATT on the path $T_3 \rightarrow T_4$ would not attract much attention of conservative readers.

[Table 5 about here]

Turning to non-cognitive traits variables, we first take up effects of parental divorce on the children's interpersonal skill development assessed by teachers [Table 5]. As with mathematics test scores, we find statistically insignificant positive pre-divorce effect. Specifically from the growth curve estimates, children of divorce appear to have showed less skilled behaviors in social relations but have enjoyed some advantage in pre-divorce period even though those findings are not replicated in other statistical models. As children of divorce started to face their fate, they tended to exhibit downgraded interpersonal skills compared to their counterparts. Borrowing the language of the data manual, children of divorce were more likely to show fall-off in "forming and maintain friendships,…, expressing feelings, ideas, and opinions in positive ways,…(Tourangeau et al. 2006)" This finding is robust with regard to choice of statistical models convincing us unfavorable realities of parental divorce. Furthermore, statistically significant path estimates on the line $T_2 \rightarrow T_4$ demonstrate that these adverse effects remained unabated even after children got through the divorce period though we failed to detect stand-alone post-divorce effects, which is confirmed by all statistical models. In addition, only unweighted sample reveals negative post-divorce effects in both OLS and matching estimators. Partly due to positive figure in the pre-divorce appraisal, however, total divorce effects fall short of statistical significance.

[Table 6 about here]

Externalizing behavior problems look like another developmental domain that was relatively unaffected by parental divorce regardless of aggregation and disaggregation of divorce stages. Apart from two scattered instances suggesting disadvantageous influence of parental divorce, there is no consistent and robust evidence supporting the traditional null hypothesis of negative effects of parental divorce on involved child(ren)'s externalizing behavior problems. Instead, we encounter some indication favoring selection perspective from the growth curve model as also hinted in the interpersonal skills dimension. Put another way, the intercept terms in the trajectory of growth curves show statistically significant, elevated initial level of divorced children in comparison with children in intact family notwithstanding weighted or unweighted sample one would fit the model. We also notice that those initial gaps had sustained throughout our study period neither widening nor shrinking.

[Table 7 about here]

Finally, we pay our attention to the differential development in the internalizing behavioral dimension between two treatment groups. There are conflicting signs in estimates on the pre-divorce effect across statistical models and sample choice but they all agree in that those point estimates are imprecisely gauged in the statistical sense. Negative consequences of parental divorce are not more pronounced in any other stages than in the then-divorce period. Statistically significant point estimates in the neighborhood of a quarter of population standard deviations (See Table 2) unfailingly demonstrate conventional consensus on adverse impacts of parental divorce on children's development at least in the internalizing problem behaviors. To be more specific, children of divorce were more likely to struggle with "anxiety, loneliness, low self-esteem, and sadness" when their parents are on stage of divorce compared to otherwise counterfactual status (Tourangeau et al. 2006). Assessment on the following stage imparts a sense that those negative consequences do not seem to have disappeared or have been exacerbated. A little reduced magnitude on the estimates of the path $T_2 \rightarrow T_4$ and diversion of statistical significance may be suggestive, and only suggestive, for weak resilience on the population level. We witness that total divorce effects are not strong enough to reject the null hypothesis whether we conceive it by the path $T_1 \rightarrow T_2 \rightarrow T_4$ or by the route $T_1 \rightarrow T_2 \rightarrow T_3 \rightarrow T_4$.

Additional observations beyond domain-specific areas are worth discussing to more detailed extent in this section. We failed to uncover 1) any pre-divorce effect, 2) resilience parameters on the population level, 3), more or less related to them, total divorce effects. Regarding the pre-divorce effect, several explanations can be come up with. First, two years of then-divorce period might be wide enough to include initiation and development of marital strains and discord so that some portion of the pre-divorce effect has been assigned to the then-divorce effect. This interpretation is congruent with our results pointing toward statistically significant then-divorce effect. However, our descriptive statistics on marital happiness measured at T_1 do not corroborate this point (See Table A in Appendix). Conversely, one-year window of pre-divorce period might be too short to capture the pre-divorce effect not only because children might already have got accustomed to unfavorable daily lives but also because only negligible change can take place in such a short period even though exogenous shock is substantial. Another possibility includes that as Amato argued, not all divorce might be characterized by preceding marital conflicts before related couple filed an annulment or decided to separate (Amato 2002; Amato & Hohmann-Marriott 2007). Or considerate parents doomed to divorce might have tried to conceal or at least suppress their emotional outburst in front of their children.

Effect parameters for the post-divorce impact or resilience as defined in our theoretical discussion along either the line of $T_3 \rightarrow T_4$ or the line of $(T_2 \rightarrow T_4) - (T_2 \rightarrow T_3)$ have not passed their statistical tests except some local instances. Quite contrary to the resilience hypothesis, we detected negative post-divorce effect for mathematics test score. To catch up recent development in resilience literature and isolate comparatively more resilient subpopulation requires additional works beyond this paper. It is most compatible with our results to conclude that there is no compelling evidence to support resilience argument on the population level. Imprecise estimates on both the pre-divorce effect and the post-divorce effect deem responsible for the statistically insignificant total divorce effects. To avoid impression that statistically insignificant total divorce effects may suggest rejection of the traditional null hypothesis of negative divorce effects, we stress that discovery on the combined effects of the then- and post-divorce stage in important developmental domains constitutes one of our major contribution to the literature. Further, absence of the post-divorce effect combined with presence of negative then-effect suggests continuing, if not aggravating, developmental gap between children of divorce and those with continuously married biological parents.

7. Discussion and conclusion

So far, we have examined parental divorce effects on several developmental domains of engaged children after constructing three analytically distinct divorce stages: pre-, then-, and post-divorce period. Under guidance of statistically sophisticated methods, we have shown that effects of parental divorce are stage-specific in addition to domain-specific. To summarize our findings, 1) children of divorce lagged behind in mathematics test scores during and after they underwent parental divorce (significant combined effects of the then- and post-divorce effects), 2) as to the interpersonal skills, we detect negative then-divorce effect and negative combined effects of the then- and post-divorce effects, 3) we found pronounced then-divorce effect in the internalizing behavioral dimension, 4) we failed to ascertain presumably negative consequences of parental divorce in reading test scores and the externalizing behavioral problems in any stage of the time line. Additionally, we failed to detect statistically significant estimates on 1) the pre-divorce effect, 2) resilience parameters on the population level, 3), more or less related to them, total divorce effects.

Discussion on several limitations of this paper is in order to properly evaluate our contribution as well as to set out our future task for more sophisticated and meaningful research program. We are mostly concerned with measurement errors in our non-cognitive trait scales not only because they were measured as an average of several individual items but also because they were assessed by different teachers across longitudinal survey waves. Data collectors also cautioned against building longitudinal models with non-cognitive traits as response variables (Tourangeau et al. 2006). As such, our estimates in this paper should be considered most plausible guess at the current stage rather than definitive works for or against specified hypotheses.

Because we traced children only two years after parental divorce occurred, one may be concerned with short time span such that either latent negative effect or resilience effect could not be fully accounted for. As Cherlin commented in his well-received introductory book, effects of parental divorce may be latent in a sense that devastating results may be fully realized only after children of divorce grow up (Cherlin 2008; Wallerstein & Lewis 2004). On the opposite side, agreeing with the point that negative then-divorce effect may reflect true reality, some scholars would maintain that most children recover as time passes by. Even though our current analyses do not support either of such views, fortunately, ECLS-K is an ongoing survey and 8th grade wave are supposed to be released soon, opening another opportunity to rigorously validate or invalidate debated theoretical positions.

Our work presented in this paper is confined to those children who experienced parental divorce during the period of the spring first grade and third grade or in their 7-9 to 9-11 years old in terms of age. This limitation means, among other things, that results reported here may not apply to, for example, those children who experience parental divorce in adolescence.

Supposedly children in early childhood are relatively less attuned to environmental changes in particular involving emotional reconfiguration than those in adolescence (Papalia et al. 2004). This observation alone resists unwarranted generalization of our results and calls for extension of our analytical framework for better understanding on divorce and development of involved children.

References

Abadie, Alberto and Guido W. Imbens. 2002. “Simple and Bias-Corrected Matching Estimators.” Technical report, Department of Economics, University of California, Berkeley. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=341300 as of February, 2010.

Abadie, Alberto and Guido W. Imbens. 2006. “Large Sample Properties of Matching Estimators for Average Treatment Effects.” Econometrica 74(1):235-267.

Amato, Paul R. 1993. “Children’s Adjustment to Divorce: Theories, Hypotheses, and Empirical Support.” *Journal of Marriage and the Family* 55:23-38.

_____. 2000. “The Consequences of Divorce for Adults and Children.” *Journal of Marriage and Family* 62(4):1269-1287.

_____. 2001. “Children of Divorce in the 1990s: An Update of the Amato and Keith (1991) Meta-Analysis.” *Journal of Family Psychology* 15(3):355-370.

_____. 2002. “Good Enough Marriages: Parental Discord, Divorce, and Children’s Well-being.” *Virginia Journal of Social Policy & the Law* 9:71-94.

_____. 2003. “Reconciling Divergent Perspectives: Judith Wallerstein, Qunatitative Family Research, and Children of Divorce.” *Family Relations* 52(4):332-339.

Amato, Paul R. and Alan Booth. 1997. *A Generation at Risk: Growing Up in an Era of Family Upheaval.* Cambridge, MA: Harvard University Press.

Amato, Paul R. and Bruce Keith. 1991. “Parental Divorce and Adult Well-being: A Meta-Analysis.” *Journal of Marriage and Family* 53(1):43-58.

Amato, Paul R. and Bryndl Hohmann-Marriott. 2007. “A Comparison of High- and Low-Distress Marriages that End in Divorce.” *Journal of Marriage and Family* 69:621-638.

Astone, Nan Marie and Sara S. McLanaha. 1994. “Family Structure, Residential Mobility, and School Dropout: A Research Note.” *Demography* 31(4):575-584.

Aughinbaugh, Alison, Charles R. Pierret, and Donna S. Rothstein. 2005. “The Impact of Family Structure Transitions on Youth Achievement: Evidence from the Children of the NLSY79.” *Demography* 42(3):447-468.

Bates, Douglas. 2008. “Fitting Mixed-Effects Models Using the lme4 Package in R.” *International Meeting of the Psychometric Society*. Retrieved from http://www.stat.wisc.edu/~bates/IMPS2008/lme4-4a4.pdf as of February, 2010.

Boyle, Paul J., Hill Kulu, Thomas Cooke, Cernon Gayle, and Clara H. Mulder. 2008. “Moving and Union Dissolution.” *Demography* 45(1):209-222.

Bulanda, Jennifer Roebuck and Susan L. Brown. 2007. “Race-ethnic Differences in Marital Quality and Divorce.” *Social Science Research* 36:945-967.

Carlson, Marcia, Sara McLanahan, and Paula England. 2004. “Union Formation in Fragile Families.” *Demography* 41(2):237-261.

Cherlin, Andrew J. 2008. *Public & Private Families: An Introduction*. McGraw-Hill.

Cherlin, Andrew J., Frank F. Furstenberg, Jr., P. Lindsay Chase-Lansdale, Kathleen E. Kiernan, Philip K. Robins, Donna Ruane Morrison, and Julien O. Teitler. 1991. “Longitudinal Studies of Effects of Divorce on Children in Great Britain and the United States.” *Science* 252(5011):1386-1389.

Cherlin, Andrew J., P. Lindsay Chase-Lansdale, and C. McRae. 1998. “Effects of Divorce on Mental Health throughout the Life Course.” *American Sociological Review* 63:239-249.

Cooper, Carey E., Sara S. McLanahan, Sarah O. Meadows, and Jeanne Brooks-Gunn. 2009. "Family Structure Transitions and Maternal Parenting Stress." *Journal of Marriage and Family* 71(3):558-574.

Cui, Ming, M. Brent Donnellan, and Rand D. Conger. 2007. "Reciprocal Influences between Parents' Marital Problems and Adolescent Internalizing and Externalizing Behavior." *Developmental Psychology* 43(6):1544-1552.

Dehejia, Rajeev H. and Sadek Wahba. 2002. "Propensity Score-Matching Methods for Nonexperimental Causal Studies." *Review of Economics and Statistics* 84(1):151-161.

Diamond, Alexis and Jasjeet S. Sekhon. 2006. "Genetic Matching for Estimating Causal Effects: A General Multivariate Matching Method for Achieving Balance in Observational Studies." UC Berkeley: Institute of Governmental Studies. Retrieved from: http://escholarship.org/uc/item/8gx4v5qt

Efron, Bradley and Robert J. Tibshirani. 1993. *An Introduction to the Bootstrap*. New York, NY: Chapman & Hall.

Emery, Robert E. 1982. "Interparental Conflict and the Children of Discord and Divorce." *Psychological Bulletin* 92(2):310-330.

Elwert, Felix and Christopher Winship. 2008. "Endogenous Selection Bias." Unpublished manuscript.

Fomby, Paula and Andrew J. Cherlin. 2007. "Family Instability and Child Well-Being." American Sociological Review 72:181-204.

Gautier, Pieter A., Michael Svarer, and Coen N. Teulings. 2009. "Sin City? Why is the Divorce Rate Higher in Urban Areas?" *Scandinavian Journal of Economics* 111(3):439-456.

Glenn, Norval D. and Beth Ann Shelton. 1985. "Regional Differences in Divorce in the United States." *Journal of Marriage and Family* 47(3):641-652.

Heckman, James J. Hidehiko Ichimura, Petra Todd. 1998. "Matching as an Econometric Evalutation Estimator." *Review of Economic Studies* 65(2):261-294.

Heckman, James J. and Salvador Navarro-Lozano. 2004. "Using Matching, Instrumental Variables, and Control Functions to Estimate Economic Choice Models." *Review of Economics and Statistics* 86(1):30-57.

Hetherington, E. Mavis. 1979. "Divorce: A Child's Perspective." *American Psychologist* 34(10):851-858.

_____. 2003. "Social Support and the Adjustment of Children in Divorced and Remarried Families." *Childhood* 10(2):217-236.

_____. 2005. "Divorce and the Adjustment of Children." *Pediatrics in Review* 26:163-169.

Hirano, Keisuke, Guido W. Imbens, and Geert Ridder. 2003. "Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score." *Econometrica* 71(4):1161-1189.

Jenkins, Jennifer, Anna Simpson, Judy Dunn, Jon Rasbash, and Thomas G. O'Connor. 2005. "Mutual Influence of Marital Conflict and Children's Behavior Problems: Shared and Nonshared Family Risks." *Child Development* 76(1):24-39.

Kaplan, Howard B. 2005. "Understanding the Concept of Resilience." in *Handbook of Resilience in Children* edited by Sam Goldstein and Robert B. Brooks. 39-47. New York, NY: Kluwer Academic/Plenum Publishers.

Kelly, Joan B. and Robert E. Emery. 2003. "Children's Adjustment Following Divorce: Risk and Resilience Perspectives." *Family Relations* 52(4):352-362.

Koerner, Susan Silverberg, Den Yelle Baete Kenyon, Lela A. Rankin. 2006. "Growing Up Faster? Post-Divorce Catalysts in the Mother-Adolescent Relationship." *Journal of Divorce & Remarriage* 45(3):25-41.

Lansford, Jennifer E. 2009. "Parental Divorce and Children's Adjustment." Perspectives on Psychological Science 4(2):140-152.

Lohr, Sharon L. 1999. *Sampling: Design and Analysis*. Pacific Grove, CA: Duxbury Press.

McLanahan, Sara. S. and Gary Sandefur. 1994. *Growing up with a Single Parent: What Hurts, What Helps?*. Cambridge, MA: Harvard University Press.

Morrison, Donna Ruane and Andrew J. Cherlin. 1995. "The Divorce Process and Young Children's Well-Being: A Prospective Analysis." *Journal of Marriage and Family* 57(3):800-812.

Papalia, Diane E., Sally Wendkos Olds, and Ruth Duskin Feldman. 2004. *A Child's World: Infancy through Adolescence, Updated Ninth Edition*. New York, NY: McGraw-Hill.

Petersen, Richard R. 1996. "A Re-Evaluation of the Economic Consequences of Divorce." *American Sociological Review* 61(3):528-536.

Raudenbush, Stephen W. and Anthony S. Bryk. 2002. *Hierarchical Linear Models: Applications and Data Analysis Methods*. Thousand Oaks: Sage Publications.

Rosenbaum, Paul R. and Donald B. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70(1):41-55.

Rubin, Donald B. 1973. "Matching to Remove Bias in Observational Studies." *Biometrics* 29:159-183.

Sekhon, Jasjeet S. Forthcoming. "Multivariate and Propensity Score Matching Software with Automated Balance Optimization: The Matching package for R." *Journal of Statistical Software*.

Seltzer, Judith A. 1994. "Consequences of Marital Dissolution for Children." *Annual Review of Sociology* 20:235-266.

Schermerhorn, Alice C., E. Mark Cummings, Catherine A. DeCarlo, and Patrick T. Davies. 2007. "Children's Influence in the Marital Relationship." *Journal of Family Psychology* 21(2):259-269.

Sigle-Rushton, Wendy, John Hobcraft, and Kathleen Kiernan. 2005. "Parental Divorce and Subsequent Disadvantage: a Cross-Cohort Comparison." *Demography* 42(3):427-446

Singer, Judith D. and John B. Willett. 2003. Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence. New York, NY: Oxford University Press.

Smith, Herbert L. 1997. "Matching with Multiple Controls to Estimate Treatment Effects in Observational Studies." *Sociological Methodology* 27:325-353.

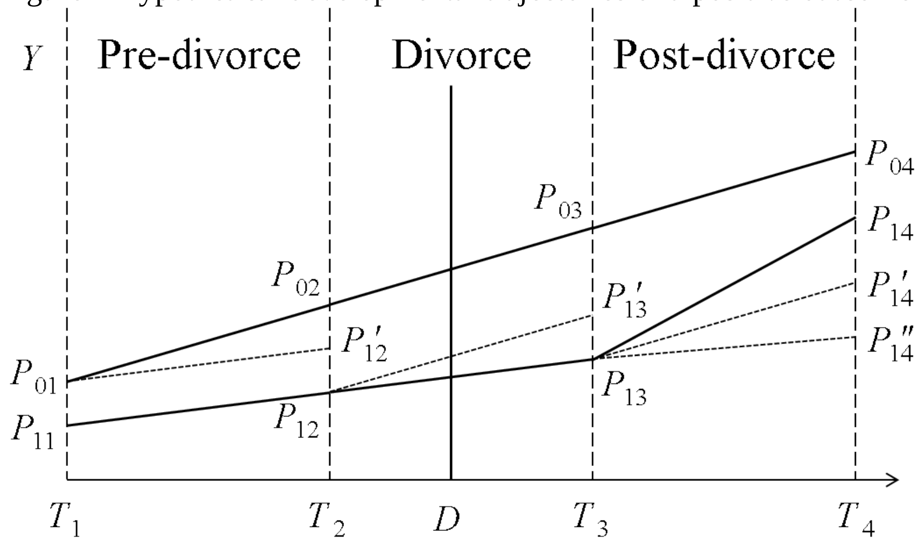
Smock, Pamela J. 1994. "Gender and the Short-Run Economic Consequences of Marital Disruption." *Social Forces* 73(1):243-62

Thomson, Elizabeth, Jane Mosley, Thomas L. Hanson, and Sara S. McLanahan. 2001. "Remarriage, Cohabitation, and Changes in Mothering Behavior." *Journal of Marriage and Family* 63(2):370-380.

Tourangeau, K., C. Nord, Lê T., J.M. Pollack, and S. Atkins-Burnett. 2006. *Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K), Combined User's Manual for the ECLS-K Fifth-Grade Data Files and Electronic Codebooks (NCES 2006-032).* U.S. Department of Education. Washington, DC: National Center for Education Statistics.

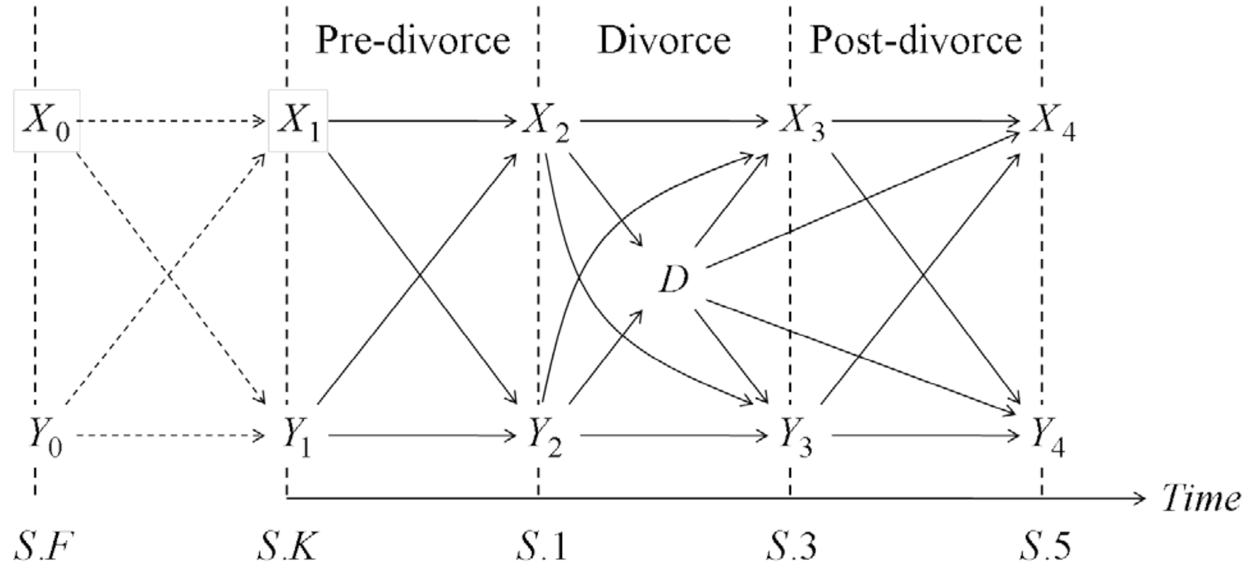
Wallerstein, Judith S. and Julia M. Lewis. 2004. “The Unexpected Legacy of Divorce: Report of a 25-Year Study.” *Psychoanalytic Psychology* 21(3):353-370.
Weiss, Robert S. 1979. “Growing Up a Little Faster: The Experience of Growing Up in a Single-Parent Household.” *Journal of Social Issues* 35(4):97-111.
Westat. 2007. *WesVar 4.3 User’s Guide*. Rockville,MD: Westat

Figure 1 Hypothetical developmental trajectories of a positive outcome



Note: Notations in points are constructed such that first subscript refers to the treatment status (zero denotes child in an intact family and unity child of divorce) and second subscript time. Prime and double primes are introduced for theoretically different scenarios from those hypothesized in this paper.

Figure 2 Causal directed acyclic graph



Note: S.F and S.K refer to fall and spring of Kindergarten respectively and S.1, S.3, and S.5 index spring 1st grade, 3rd grade, and 5th grade respectively. D denotes divorce, $X_t : t \in \{0,1,2,3,4\}$ a set of covariates, and Y_t a outcome variable. T_0 is necessary to provide confounding variables for T_1 outcomes in piece-wise growth curve models but do not contain parameters of interest so that we brush those relations with dotted lines.

Table 1 Specification for piecewise growth curve

	T_0	T_1	T_2	T_3
a_1	0	1	1	1
a_2	0	0	1	1
a_3	0	0	0	1

Table 2 Descriptive statistics for treatment, outcome and lagged outcome variables

Var. name	Description	T	Unweighted: Intact (=0) F/M[1)]	P/S[2)]	Mn[3)]	Mx[4)]	Unweighted: Divorce(=1) F/M[1)]	P/S[2)]	Mn[3)]	Mx[4)]	Weighted ("c1_6fp0"): Intact F/M[1)]	P/S[2)]	Weighted ("c1_6fp0"): Divorce F/M[1)]	P/S[2)]	Original variables
Divorce			3,447	96.0			142	4.0				93.9		6.1	See note.
math0	Mathematics test score	T0	26.5	9.3	8.2	84.4	22.8	7.6	10.4	61.5	25.8	9.2	22.4	6.3	c"t"r3mscl for "t"=1,2,4,5,6
math1		T1	37.6	11.4	11.9	102.6	33.8	10.6	16.2	80.6	36.6	11.2	31.6	9.3	
math2		T2	63.7	16.1	17.8	120.5	59.1	15.3	24.1	96.6	62.4	15.9	58.7	14.5	
math3		T3	99.9	19.0	36.9	146.6	93.7	18.9	47.6	137.6	98.5	19.7	91.7	17.3	
math4		T4	121.2	17.4	51.6	150.9	114.0	18.6	55.5	145.9	120.1	18.3	109.6	18.6	
read0	Reading test score	T0	32.2	10.5	15.5	118.5	28.5	7.0	16.6	48.5	31.8	10.4	28.4	7.0	c"t"r3rscl for "t"=1,2,4,5,6
read1		T1	44.5	14.0	17.5	128.1	39.7	9.1	22.2	73.5	43.9	14.1	39.2	8.2	
read2		T2	78.7	21.1	22.6	163.1	71.4	18.4	35.2	124.8	77.6	21.0	72.0	17.9	
read3		T3	127.8	21.2	49.9	176.9	118.1	22.8	58.6	171.4	126.0	22.2	120.4	22.7	
read4		T4	148.1	19.0	61.4	181.2	139.2	19.9	64.7	180.4	146.5	20.2	138.2	18.6	
interp0	Teacher-assessed interpersonal skills	T0	2.1	0.6	0.3	3.0	2.0	0.6	0.6	3.0	2.1	0.6	2.0	0.7	t"t"interp for "t"=1,2,4,5,6
interp1		T1	2.3	0.6	0.0	3.0	2.1	0.6	0.8	3.0	2.3	0.6	2.1	0.6	
interp2		T2	2.2	0.6	0.2	3.0	2.2	0.6	0.8	3.0	2.2	0.6	2.2	0.7	
interp3		T3	2.2	0.6	0.0	3.0	2.0	0.6	0.0	3.0	2.2	0.6	2.0	0.6	
interp4		T4	2.2	0.6	0.0	3.0	2.0	0.7	0.2	3.0	2.2	0.6	1.9	0.7	
extern0	Teacher - assessed externalizing behavior problems	T0	0.5	0.5	0.0	3.0	0.6	0.6	0.0	2.6	0.5	0.6	0.7	0.7	t"t"extern for "t"=1,2,4,5,6
extern1		T1	0.5	0.5	0.0	3.0	0.7	0.6	0.0	3.0	0.5	0.5	0.7	0.6	
extern2		T2	0.5	0.5	0.0	3.0	0.7	0.6	0.0	2.8	0.5	0.6	0.7	0.6	
extern3		T3	0.6	0.5	0.0	2.8	0.7	0.6	0.0	3.0	0.6	0.5	0.8	0.7	
extern4		T4	0.5	0.5	0.0	3.0	0.7	0.6	0.0	2.8	0.6	0.5	0.7	0.7	
intern0	Teacher-assessed internalizing behavior problems	T0	0.5	0.5	0.0	3.0	0.5	0.5	0.0	2.3	0.5	0.5	0.6	0.5	t"t"intern for "t"=1,2,4,5,6
intern1		T1	0.5	0.4	0.0	3.0	0.6	0.5	0.0	2.3	0.5	0.5	0.6	0.6	
intern2		T2	0.5	0.5	0.0	3.0	0.6	0.4	0.0	2.3	0.5	0.5	0.5	0.5	
intern3		T3	0.5	0.5	0.0	3.0	0.8	0.6	0.0	2.5	0.6	0.5	0.8	0.6	
intern3		T4	0.6	0.5	0.0	3.0	0.7	0.6	0.0	2.7	0.6	0.5	0.7	0.7	

Note: To construct divorce variable, "p4hparnt, p5hparnt, p4curmar, p5curmar" were used. 1) Frequency/mean, 2) percentage/population standard deviation, 3) minimum, 4) maximum. 0,1,2,3, and 4 in time and variable names means that those variables were measured at fall and spring of kindergarten and spring of first, third, and fifth grade. We do not show weighted frequency because it is irrelevant.

Table 3. Estimates from statistical models with mathematics test score as a response variable

Model	$T_0 \rightarrow T_1$	Pre-effect $T_1 \rightarrow T_2$	Then-effect $T_2 \rightarrow T_3$	Post-effect: resilience $T_3 \rightarrow T_4$	Post-effect: resilience $(T_2 \rightarrow T_4) -(T_0 \rightarrow T_1)$	Then- plus post- effects $T_2 \rightarrow T_3 \rightarrow T_4$	Then- plus post- effects $T_2 \rightarrow T_4$	Total effects $T_1 \rightarrow T_2 \rightarrow T_3 \rightarrow T_4$	Total effects $T_1 \rightarrow T_2 \rightarrow T_4$
Column name	1	2	3	4	5 (=7-4)	6 (=3+4)	7	8 (=2+3+4)	9 (=2+7)
OLS									
Unweighted.		0.115	-0.583	-1.101	-1.385	-1.684	-1.968	-1.569	-1.853
Asymp. SE		(0.935)	(1.038)	(0.780)			(1.046) †		
Boot. SE		(0.939)	(0.943)	(0.796)	(0.819) †	(1.241)	(1.036) †	(1.530)	(1.366)
Weighted		2.485 *	-1.791	-3.181 *	-3.090 **	-4.972 *	-4.881 **	-2.487	-2.396
P.J.K. SE		(1.214)	(1.358)	(1.230)	(1.079)	(2.038)	(1.785)	(2.213)	(1.980)
Matching									
Unweighted		1.289	-0.566	-2.621	-3.119	-3.186	-3.685	-1.898	-2.396
Asymp. SE		(1.377)	(1.426)	(1.864)			(1.544) *		
Boot. SE		(1.812)	(1.502)	(1.760)	(1.979)	(2.067)	(1.692) *	(2.616)	(1.940)
Weighted		-0.378	-2.633	-7.526 *	-5.561 ***	-10.159 *	-8.194 **	-10.537 *	-8.572 **
P.J.K. SE		(1.671)	(2.064)	(3.729)	(1.630)	(4.764)	(2.435)	(5.142)	(3.122)
Growth curve									
Unweighted	0.295	0.007	-0.575	-0.821		-1.396		-1.389	
Asympt. SE	(0.575)	(1.243)	(1.703)	(1.436)		(1.223)		(0.949)	
Weighted	0.191	0.818	-1.436	-0.432		-1.868 *		-1.051	
P.J.K SE (90)	(0.396)	(0.773)	(0.896)	(1.034)		(0.934)		(0.880)	
Log unweighted	N.A.								
Log weighted	0.864 **	-0.232	-1.367	0.064		-1.302		-1.534 †	
P.J.K SE (90)	(0.314)	(0.947)	(1.129)	(1.341)		(1.256)		(0.914)	

Note: †<0.1; *<0.05; *<0.01; ***<0.001. Standard errors in parenthesis. For unweighted OLS and matching estimates, we mark p-value on standard error because point estimates are the same for the two different standard errors. Asymp. SE denotes asymptotic standard error, Boot. SE bootstrapped standard errors, P.J.K. SE paired Jackknife standard errors. In growth curve models, “log” means all non-cognitive trait variables are logged after adding unity. N.A. means that estimates are not available due to convergence problem. Number in parenthesis after SE means how many replicate weights were used due to convergence problem.

Table 4. Estimates from statistical models with reading test score as a response variable

Model		Pre-effect	Then-effect	Post-effect: resilience		Then- plus post- effects		Total effects	
	$T_0 \to T_1$	$T_1 \to T_2$	$T_2 \to T_3$	$T_3 \to T_4$	$(T_2 \to T_4) -(T_0 \to T_1)$	$T_2 \to T_3 \to T_4$	$T_2 \to T_4$	$T_1 \to T_2 \to T_3 \to T_4$	$T_1 \to T_2 \to T_4$
Column name	1	2	3	4	5 (=7-4)	6 (=3+4)	7	8 (=2+3+4)	9 (=2+7)
OLS									
Unweighted.		-0.757	-2.824	-0.933	0.160	-3.757	-2.664	-4.514	-3.421
Asymp. SE		(1.179)	(1.267)*	(0.932)			(1.204)*		
Boot. SE		(1.092)	(1.418)*	(1.032)	(1.261)	(1.532)*	(1.213)*	(2.104)*	(1.777)†
Weighted		1.258	0.015	-1.964	-2.091	-1.949	-2.076	-0.692	-0.819
P.J.K. SE		(2.054)	(2.781)	(1.208)	(1.898)	(2.304)	(1.559)	(3.369)	(2.628)
Matching									
Unweighted		-0.799	-1.619	-1.880	0.226	-3.500	-1.393	-4.299	-2.192
Asymp. SE		(1.678)	(2.004)	(1.610)			(1.953)		
Boot. SE		(1.706)	(2.754)	(1.556)	(1.581)	(3.058)	(2.470)	(3.766)	(3.203)
Weighted		-0.736	-1.430	-5.790**	-2.373	-7.219†	-3.802	-7.956	-4.539
P.J.K. SE		(3.116)	(4.219)	(1.967)	(2.642)	(4.258)	(3.220)	(6.695)	(5.788)
Growth curve									
Unweighted	-0.325	-0.293	-1.935	1.889		-0.046		-0.338	
Asympt. SE	(0.668)	(1.404)	(1.800)	(1.695)		(1.469)		(1.144)	
Weighted	N.A.								
Log unweighted	N.A.								
Log weighted	0.104	-1.124**	-1.736*	1.648**		-0.088		-1.212**	
P.J.K SE (14)	(0.144)	(0.300)	(0.686)	(0.420)		(0.405)		(0.302)	

Note: †<0.1; *<0.05; *<0.01; ***<0.001. Standard errors in parenthesis. For unweighted OLS and matching estimates, we mark p-value on standard error because point estimates are the same for the two different standard errors. Asymp. SE denotes asymptotic standard error, Boot. SE bootstrapped standard errors, P.J.K. SE paired Jackknife standard errors. In growth curve models, "log" means all non-cognitive trait variables are logged after adding unity. N.A. means that estimates are not available due to convergence problem. Number in parenthesis after SE means how many replicate weights were used due to convergence problem.

Table 5. Estimates from statistical models with interpersonal social skill as a response variable

Model	$T_0 \to T_1$	Pre-effect $T_1 \to T_2$	Then-effect $T_2 \to T_3$	Post-effect: resilience $T_3 \to T_4$	Post-effect: resilience $(T_2 \to T_4) - (T_0 \to T_1)$	Then- plus post- effects $T_2 \to T_3 \to T_4$	Then- plus post- effects $T_2 \to T_4$	Total effects $T_1 \to T_2 \to T_3 \to T_4$	Total effects $T_1 \to T_2 \to T_4$
Column name	1	2	3	4	5 (=7-4)	6 (=3+4)	7	8 (=2+3+4)	9 (=2+7)
OLS									
Unweighted.		0.031	-0.104	-0.103	-0.030	-0.207	-0.134	-0.176	-0.103
Asymp. SE		(0.047)	(0.047) *	(0.047) *			(0.047) **		
Boot. SE		(0.051)	(0.046) *	(0.054) †	(0.062)	(0.072) **	(0.057) *	(0.089) *	(0.072)
Weighted		0.061	-0.137 †	-0.075	0.012	-0.211 *	-0.125	-0.150	-0.064
P.J.K. SE		(0.067)	(0.080)	(0.097)	(0.121)	(0.088)	(0.076)	(0.114)	(0.104)
Matching									
Unweighted		0.016	-0.119	-0.150	-0.074	-0.269	-0.193	-0.253	-0.177
Asymp. SE		(0.074)	(0.067) †	(0.070) *			(0.067) **		
Boot. SE		(0.073)	(0.070) †	(0.071) *	(0.092)	(0.094) **	(0.072) **	(0.118) *	(0.096) †
Weighted		0.178	-0.233 *	-0.149	0.015	-0.382 **	-0.218 *	-0.204	-0.040
P.J.K. SE		(0.132)	(0.102)	(0.130)	(0.114)	(0.130)	(0.096)	(0.201)	(0.185)
Growth curve									
Unweighted	N.A.								
Weighted	-0.107 **	0.176 **	-0.167 **	0.067		-0.099		0.077	
P.J.K. SE (88)	(0.034)	(0.059)	(0.060)	(0.061)		(0.074)		(0.052)	
Log unweighted	N.A.								
Log weighted	-0.048 *	0.080 *	-0.070 *	-0.027		-0.096 *		-0.017	
P.J.K. SE (89)	(0.023)	(0.039)	(0.032)	(0.030)		(0.037)		(0.022)	

Note: †<0.1; *<0.05; *<0.01; ***<0.001. Standard errors in parenthesis. For unweighted OLS and matching estimates, we mark p-value on standard error because point estimates are the same for the two different standard errors. Asymp. SE denotes asymptotic standard error, Boot. SE bootstrapped standard errors, P.J.K. SE paired Jackknife standard errors. In growth curve models, “log” means all non-cognitive trait variables are logged after adding unity. N.A. means that estimates are not available due to convergence problem. Number in parenthesis after SE means how many replicate weights were used due to convergence problem.

Table 6. Estimates from statistical models with externalizing problem behavior as a response variable

Model		Pre-effect	Then-effect	Post-effect: resilience		Then- plus post- effects		Total effects	
	$T_0 \to T_1$	$T_1 \to T_2$	$T_2 \to T_3$	$T_3 \to T_4$	$(T_2 \to T_4)$ $-(T_0 \to T_1)$	$T_2 \to T_3$ $\to T_4$	$T_2 \to T_4$	$T_1 \to T_2$ $\to T_3 \to T_4$	$T_1 \to T_2$ $\to T_4$
Column name	1	2	3	4	5 (=7-4)	6 (=3+4)	7	8 (=2+3+4)	9 (=2+7)
OLS									
Unweighted.		0.051	0.053	0.044	0.000	0.096	0.053	0.148	0.104
Asymp. SE		(0.040)	(0.039)	(0.038)			(0.039)		
Boot. SE		(0.041)	(0.043)	(0.047)	(0.055)	(0.060)	(0.048)	(0.082) †	(0.071)
Weighted		0.046	0.129	-0.001	-0.088	0.128	0.041	0.174	0.087
P.J.K. SE		(0.043)	(0.113)	(0.104)	(0.134)	(0.123)	(0.093)	(0.125)	(0.101)
Matching									
Unweighted		0.055	0.021	0.064	0.091	0.086	0.113	0.140	0.167
Asymp. SE		(0.059)	(0.062)	(0.062)			(0.062) †		
Boot. SE		(0.067)	(0.073)	(0.073)	(0.100)	(0.112)	(0.066) †	(0.155)	(0.088) †
Weighted		-0.089	0.146	0.069	-0.037	0.215	0.109	0.126	0.020
P.J.K. SE		(0.113)	(0.111)	(0.116)	(0.151)	(0.169)	(0.131)	(0.184)	(0.182)
Growth curve									
Unweighted	0.055 †	-0.022	-0.018	-0.001		-0.019		-0.041	
Asymp. SE	(0.033)	(0.060)	(0.071)	(0.070)		(0.059)		(0.051)	
Weighted	0.071 **	-0.044	-0.020	0.022		0.002		-0.043	
P.J.K. SE (60)	(0.023)	(0.045)	(0.050)	(0.053)		(0.032)		(0.048)	
Log unweighted	N.A.								
Log weighted	0.081	-0.017	0.030	-0.069		-0.040		-0.057	
P.J.K. SE (88)	(0.057)	(0.078)	(0.072)	(0.074)		(0.066)		(0.065)	

Note: †<0.1; *<0.05; *<0.01; ***<0.001. Standard errors in parenthesis. For unweighted OLS and matching estimates, we mark p-value on standard error because point estimates are the same for the two different standard errors. Asymp. SE denotes asymptotic standard error, Boot. SE bootstrapped standard errors, P.J.K. SE paired Jackknife standard errors. In growth curve models, “log” means all non-cognitive trait variables are logged after adding unity. N.A. means that estimates are not available due to convergence problem. Number in parenthesis after SE means how many replicate weights were used due to convergence problem.

Table 7. Estimates from statistical models with internalizing problem behavior as a response variable

Model	$T_0 \to T_1$	Pre-effect $T_1 \to T_2$	Then-effect $T_2 \to T_3$	Post-effect: resilience $T_3 \to T_4$	Post-effect: resilience $(T_2 \to T_4)$ $-(T_0 \to T_1)$	Then- plus post- effects $T_2 \to T_3$ $\to T_4$	Then- plus post- effects $T_2 \to T_4$	Total effects $T_1 \to T_2$ $\to T_3 \to T_4$	Total effects $T_1 \to T_2$ $\to T_4$
Column name	1	2	3	4	5 (=7-4)	6 (=3+4)	7	8 (=2+3+4)	9 (=2+7)
OLS									
Unweighted.		0.003	0.189	0.036	-0.091	0.225	0.098	0.228	0.101
Asymp. SE		(0.038)	(0.039) ***	(0.042)			(0.043) *		
Boot. SE		(0.037)	(0.059) **	(0.051)	(0.071)	(0.073) **	(0.051) †	(0.080) **	(0.067)
Weighted		-0.074	0.217 *	0.034	-0.107	0.251 *	0.110	0.177	0.036
P.J.K. SE		(0.069)	(0.088)	(0.086)	(0.117)	(0.121)	(0.084)	(0.155)	(0.130)
Matching									
Unweighted		-0.035	0.154	0.026	-0.056	0.181	0.099	0.146	0.063
Asymp. SE		(0.058)	(0.061) *	(0.066)			(0.065)		
Boot. SE		(0.076)	(0.084) †	(0.089)	(0.108)	(0.144)	(0.071)	(0.197)	(0.102)
Weighted		-0.067	0.207 †	0.131	-0.051	0.337 *	0.155 †	0.271	0.089
P.J.K. SE		(0.088)	(0.116)	(0.095)	(0.142)	(0.136)	(0.092)	(0.195)	(0.159)
Growth curve									
Unweighted	N.A.								
Weighted	N.A.								
Log unweighted	N.A.								
Log weighted	0.030	-0.001	0.200 **	-0.211 *		-0.011		-0.012	
P.J.K. SE (90)	(0.054)	(0.077)	(0.074)	(0.085)		(0.061)		(0.062)	

Note: †<0.1; *<0.05; *<0.01; ***<0.001. Standard errors in parenthesis. For unweighted OLS and matching estimates, we mark p-value on standard error because point estimates are the same for the two different standard errors. Asymp. SE denotes asymptotic standard error, Boot. SE bootstrapped standard errors, P.J.K. SE paired Jackknife standard errors. In growth curve models, "log" means all non-cognitive trait variables are logged after adding unity. N.A. means that estimates are not available due to convergence problem. Number in parenthesis after SE means how many replicate weights were used due to convergence problem.

Appendix: Table A Descriptive statistics for confounding variables

			Unweighted								Weighted ("c1_6fp0")				Original variables
			Intact (=0)				Divorce(=1)				Intact		Divorce		
Var. name	Description	Values	F/M[1)]	P/S[2)]	Mn[3)]	Mx[4)]	F/M	P/S	Mn	Mx	F/M	P/S	F/M	P/S	
Age	age in month as of June, 2000		87.3	4.2	78.0	101.0	87.3	4.1	78.0	98.0	87.2	4.2	87.4	371.6	dobmm,dobyy
Gender	Gender	0: male	1,705	49.5			75	52.8				49.4		49.9	gender
		1: female	1,742	50.5			67	47.2				50.6		50.1	
Race	race/ethnicity	0: white	2715	78.8			104	73.2				78.4		60.2	race
		1: black	116	3.4			14	9.9				5.0		19.7	
		2: Hispanic	319	9.3			12	8.5				10.3		13.2	
		3: others	297	8.6			12	8.5				6.3		6.9	
ses0(=ses1)	Socio-economic index	T0=T1	0.4	0.7	-1.8	2.8	0.1	0.6	-1.0	2.5	0.3	0.7	-0.1	0.5	wksesl
ses2		T2	0.4	0.7	-1.6	2.9	0.1	0.6	-1.0	2.7	0.3	0.7	-0.1	0.6	w1sesl
ses3		T3	0.3	0.7	-2.1	2.6	0.0	0.7	-1.2	2.1	0.2	0.7	-0.2	0.6	w3sesl
disabl0 (=disabl1)	status of disability	0: no disab.	3,039	88.2			118	83.1				87.3		84.5	p1disabl
		1: disab.	408	11.8			24	16.9				12.8		15.5	
disabl2		0	2,993	86.8			121	85.2				85.9		86.3	p4disabl
		1	454	13.2			21	14.8				14.1		13.7	
disabl3		0	2,602	75.5			101	71.1				75.5		79.7	p5disabl
		1	845	24.5			41	28.9				24.5		20.3	
Sibling0	number of siblings in household	0: 0	338	9.8			16	11.3				9.8		11.9	p1numsib
		1: 1	1,712	49.7			60	42.3				49.7		40.3	
		2: 2	985	28.6			43	30.3				29.2		37.6	
		3: 3+	412	12.0			23	16.2				11.3		10.2	
Sibling1		0	324	9.4			17	12.0				9.5		12.4	p2numsib
		1	1,703	49.4			60	42.3				49.7		40.3	
		2	990	28.7			43	30.3				29.3		37.1	
		3	430	12.5			22	15.5				11.6		10.2	
Sibling2		0	286	8.3			16	11.3				8.3		11.7	p4numsib
		1	1,694	49.1			59	41.5				48.9		40.0	
		2	1,010	29.3			46	32.4				30.6		38.4	
		3	457	13.3			21	14.8				12.2		9.9	

Note: 1) Frequency/mean, 2) percentage/population standard deviation, 3) minimum, 4) maximum. 0,1,2,3, and 4 in time and variable names means that those variables were measured at fall and spring of kindergarten and spring of first, third, and fifth grade.

Appendix: Table A continued.

Var. name	Description	Values	Unweighted: Intact (=0) F/M[1]	P/S[2]	Mn[3]	Mx[4]	Unweighted: Divorce(=1) F/M	P/S	Mn	Mx	Weighted ("c1_6fp0"): Intact F/M	P/S	Divorce F/M	P/S	Original variables
Sibling3		0:00	271	7.9			28	19.7				7.9		18.9	p5numsib
		1:01	1,660	48.2			54	38.0				47.9		37.1	
		2:02	1,042	30.2			35	24.6				31.3		33.1	
		3: 3+	474	13.8			25	17.6				12.9		10.9	
mbirth	bio-parents married at time of birth	0: No	208	6.0			19	13.4				6.8		8.9	w1momar
		1: Yes	3,239	94.0			123	86.6				93.2		91.1	
mhappy	marital happiness as of T1 interview	0: Not too happy	35	1.0			7	4.9				0.9		4.4	p2marrig
		1: Fairly happy	695	20.2			56	39.4				20.8		33.1	
		2: Very happy	2,717	78.8			79	55.6				78.4		62.5	
mpsych	mother's psychological symptoms as of T1 interview		0.4	0.4	0.0	3.0	0.4	0.4	0.0	1.8	0.4	0.4	0.5	0.4	See note.
Urban0 (=urban1)	Location type	0: City	1,112	32.3			46	32.4				28.2		45.8	kurban_r
		1: Large Town	1,406	40.8			63	44.4				45.8		41.1	
		2: Small Town	929	27.0			33	23.2				26.0		13.1	
Urban2		0	1,117	32.4			46	32.4				28.8		45.8	r4urban
		1	1,400	40.6			63	44.4				45.4		41.1	
		2	930	27.0			33	23.2				25.9		13.1	
Urban3		0	1,112	32.3			46	32.4				27.9		46.8	r5urban
		1	1,407	40.8			63	44.4				45.7		40.1	
		2	928	26.9			33	23.2				26.4		13.1	

Note: To construct "mpsych" variable, "p2bother, p2appeti, p2blue, p2kpmind, p2depres, p2effort, p2fearfl, p2restls, p2talkls, p2lonely, p2sad, p2notgo" were used. 1) Frequency/mean, 2) percentage/population standard deviation, 3) minimum, 4) maximum. 0,1,2,3, and 4 in time and variable names means that those variables were measured at fall and spring of kindergarten and spring of first, third, and fifth grade. We do not show weighted frequency because it is irrelevant.

Appendix: Table A continued.

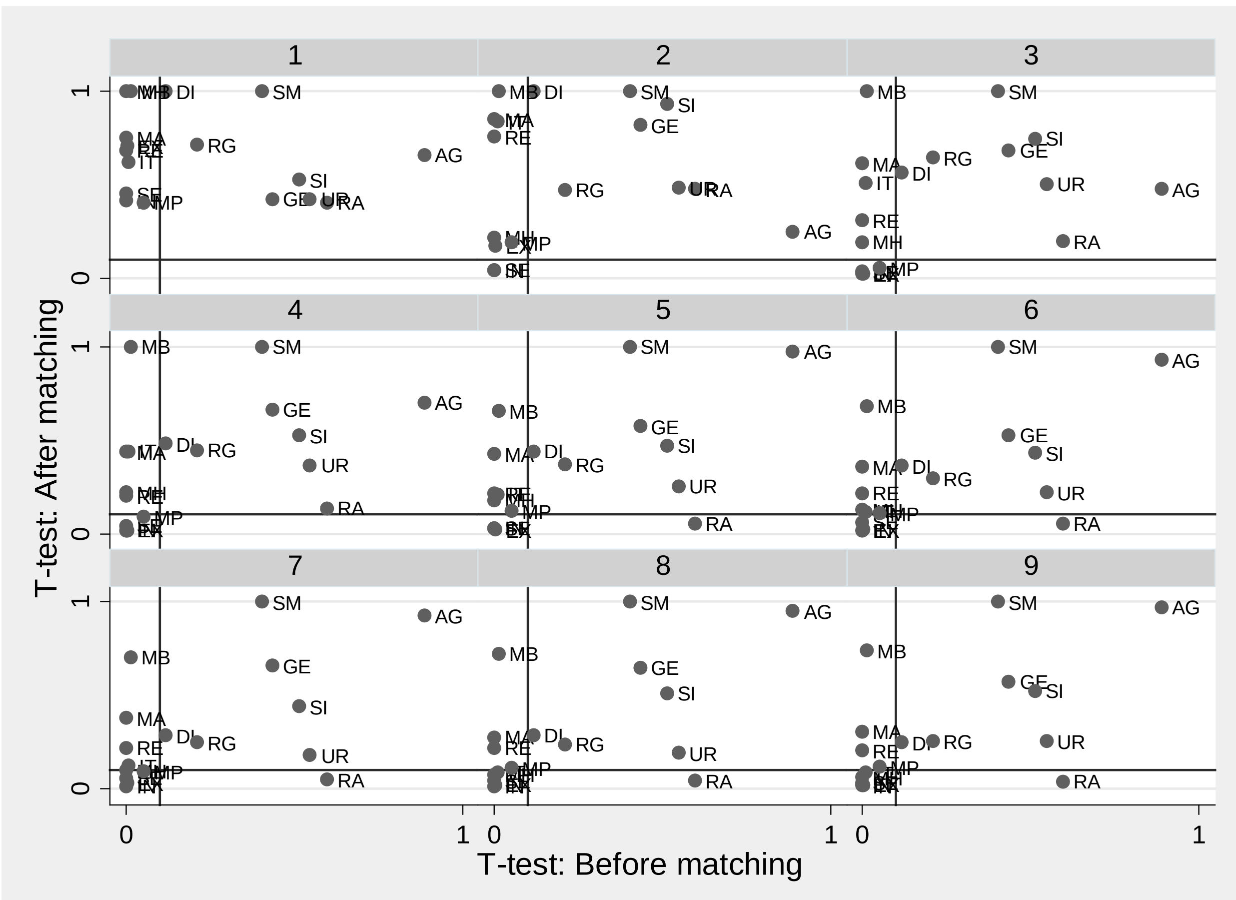
			Unweighted								Weighted ("c1_6fp0")				Original variables
			Intact (=0)				Divorce(=1)				Intact		Divorce		
Var. name	Description	Values	F/M[1)]	P/S[2)]	Mn[3)]	Mx[4)]	F/M	P/S	Mn	Mx	F/M	P/S	F/M	P/S	
Region0 (=region1)	Census region	0: North East	707	20.5			23	16.2				21.0		10.3	cregion
		1: Mid West	1,149	33.3			42	29.6				28.5		25.4	
		2: South	917	26.6			52	36.6				31.8		49.3	
		3: West	674	19.6			25	17.6				18.7		15.0	
Region2		0	707	20.5			23	16.2				21.0		10.3	r4region
		1	1,150	33.4			42	29.6				28.8		25.4	
		2	916	26.6			52	36.6				31.4		49.3	
		3	674	19.6			25	17.6				18.7		15.0	
Region3		0	706	20.5			23	16.2				21.0		10.3	r5region
		1	1,152	33.4			42	29.6				28.9		25.4	
		2	916	26.6			52	36.6				31.5		49.3	
		3	673	19.5			25	17.6				18.7		15.0	
smove0	moved school between T and T+1	0: No	3,444	99.9			142	100.0				99.7		100.0	fkchgsch
		1: Yes	3	0.1			0	0.0				0.3		0.0	
smove1		0	3,351	97.2			136	95.8				88.4		82.6	r4r2schg
		1	96	2.8			6	4.2				11.6		17.4	
smove2		0	3,186	92.4			123	86.6				86.8		80.8	r5r4schg
		1	261	7.6			19	13.4				13.2		19.2	
smove3		0	2,926	84.9			112	78.9				72.2		52.3	r6r5schg
		1	521	15.1			30	21.1				27.9		47.7	

Note: 1) Frequency/mean, 2) percentage/population standard deviation, 3) minimum, 4) maximum. 0,1,2,3, and 4 in time and variable names means that those variables were measured at fall and spring of kindergarten and spring of first, third, and fifth grade. We do not show weighted frequency because it is irrelevant.

Appendix B Report on balancing quality on matched pairs

In this appendix, we report how well the matching estimator worked in balancing values of covariates. Figure B.1 through B.3 display p-values of t-tests on the null hypothesis of no difference in covariate values between the two treatment groups before and after matching was implemented. Even though t-test does not check comprehensive dimensions of balancing, it is the most basic one to undertake to get a sense about the matching quality (Sekhon forthcoming; Dehejia &Wahba 2002). More detailed statistics such as Kolmogorov-Smirnov test provided by GenMatch are readily available on request from the author. We also insert one horizontal and vertical line pointing to 0.1 in order to aid interpretation.

Figure B.1 Covariate balancing before and after matching at T_1



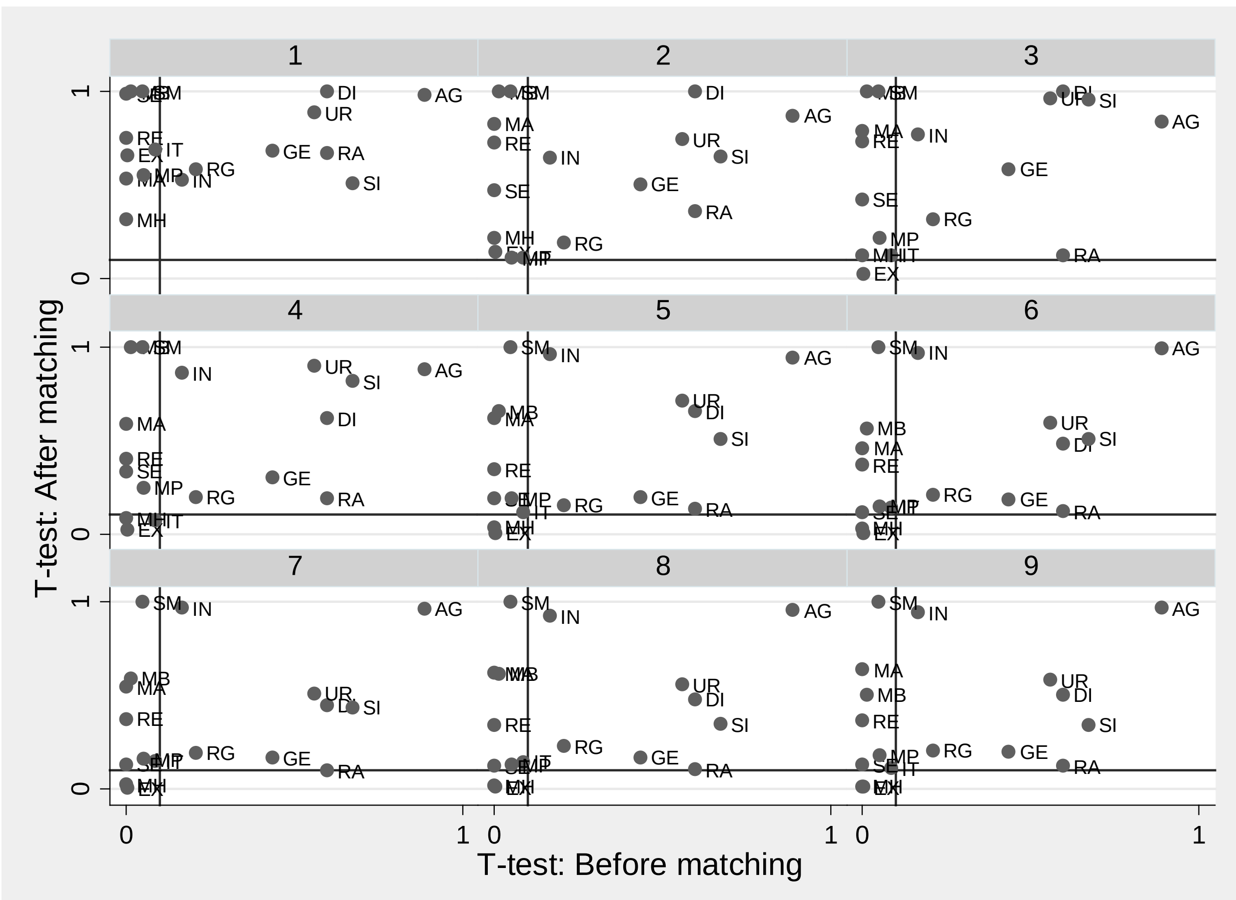
Note: Number at the head of each cell denotes number of controlled children per a treated child. Names are first two characters of variable names specified in descriptive statistics except IT(=internalizing) and RG(=region). Horizontal and vertical line refers to 0.1.

For T_1 variables, several variables were distributed unequally between two treatment groups before matching, notably mathematics and reading test scores, socio-economic index, and marital happiness reported by mother. After one-to-one matching, we find that there is no variable showing significant difference at the p-value of 0.1. As we increase the number of controlled children for a child of divorce, however, balancing quality deteriorates conspicuously.

One-to-two match gravitates the interpersonal social skills and socio-economic index into the group of significant difference and one-to-three match add the externalizing problem behaviors to the group. In general, we see that increase of controlled observations undermine one of most critical assumption in matching estimators by contaminating covariate balancing.

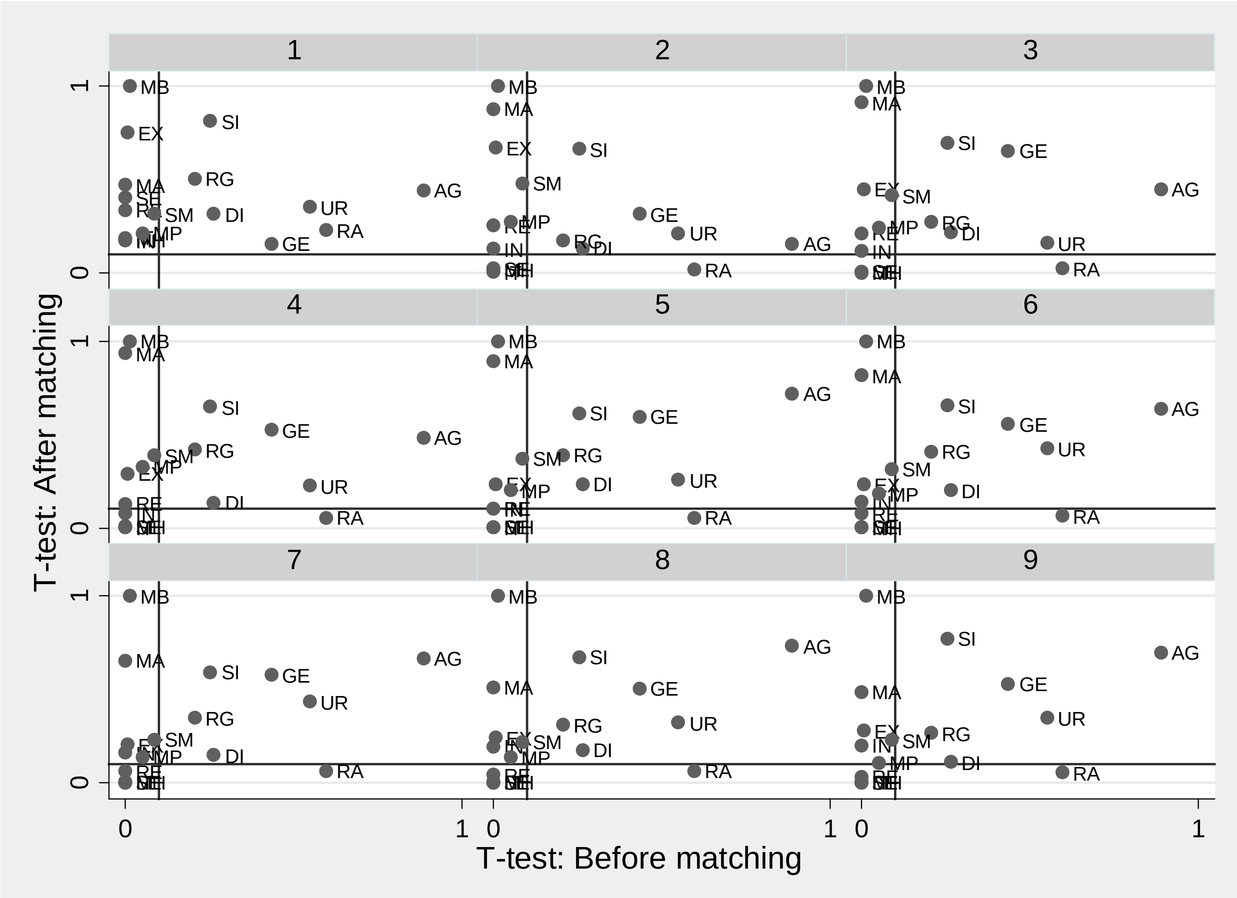
Figure B.2 and B.3 demonstrate that similar observations are applied to covariate balancing at T_2 and T_3 respectively. One interesting recognition throughout these graphs is that the most disproportionately distributed variables between the two treatment groups include mathematics and reading test score, the interpersonal social skills, the externalizing and internalizing problem behaviors, socio-economic index, marital status at birth of a child, marital happiness, and mother' psychological well-being: a predictable set of selection factors. Finally, we note that only one-to-one match generates covariate balancing across all covariates, at least for t-test.

Figure B.2 Covariate balancing before and after matching at T_2



Note: Number at the head of each cell denotes number of controlled children per a treated child. Names are first two characters of variable names specified in descriptive statistics except IT(=internalizing) and RG(=region). Horizontal and vertical line refers to 0.1.

Figure B.3 Covariate balancing before and after matching at T_3



Note: Number at the head of each cell denotes number of controlled children per a treated child. Names are first two characters of variable names specified in descriptive statistics except IT(=internalizing) and RG(=region). Horizontal and vertical line refers to 0.1.