# Parental Separations: Developments, Causes and Consequences

Raoul van Maarseveen and Yaroslav Yakymovych*

March 26, 2026

**Abstract**

Parental separation is a common event in childhood, yet its development and consequences remain poorly understood. In this paper, we use novel administrative data on residential co-location to identify all types of parental separation events among children born between 1952 and 2007 in Sweden. While high-SES households are as stable today as in the 1960s, household stability has declined sharply among low-SES households since the 1970s. Restricting attention to divorce severely understates the SES gradient and changes in parental separation over time. Turning to children's outcomes, we find little evidence that parental separations have a long-run impact on learning outcomes as measured by national test scores. In contrast, parental separations substantially reduce teacher-assigned grades, and consequently, academic track enrollment in high school. The findings suggest that school-based interventions, rather than household interventions, may be key to supporting children following parental separation.

*University of Cologne and Uppsala University. We would like to express our gratitude to Anna Bindler, Eric Chyn, Matz Dahlberg, Per-Anders Edin, Gustaf Gredeback, Ronja Helénsdotter, Martin Önnerfors, Pia Pinger, and Sebastian Siegloch, as well as seminar participants at the Uppsala Urban Lab seminar, Institute for Housing and Urban Research seminar, University of Cologne Applied seminar, and University of Nuremberg economics seminar for their helpful comments and suggestions. We would like to express our gratitude to the Institute for Housing and Urban Research at Uppsala University for providing access to administrative data for this project. Van Maarseveen gratefully acknowledges funding from the European Research Council under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. 950641)

# 1 Introduction

Despite rising rates of parental separations and single-parent households in many countries, the developments and consequences of parental separation events remain poorly understood. In this study, we use novel administrative data from Sweden to study parental separations for children born between 1952 and 2007. A key novelty is that we identify parental separations based on residential colocations on a 100x100 meter grid. As such, we do not rely on divorce to identify separation events, which as we will show, have become increasingly unrepresentative due to changing selection into marriage among low-SES households. Furthermore, the availability of data on national standardized tests and various other human capital outcomes enables us to study the effects of parental separation in more detail than previous studies.

To understand the developments in parental separations and who is affected by parental separations, we start by providing a set of four novel descriptive facts. We first focus on the changing nature of the SES gradient in parental separations. There was a stable and positive relationship between parental separation rates and SES prior to 1970. This relationship changed around 1970. Among recent cohorts, an additional year of maternal education is associated with a 4 pp decline in the parental separation probability. In contrast, children born to high-SES families today are as likely to experience a parental separation event as those born in 1960. Consequently, the entire increase in parental separations over the last five decades has been driven by a decline in family stability among low-SES households.

Second, we provide novel evidence regarding the type of parental separation events. The majority of parental separations among all groups are driven by separations that occur during childhood, rather than single parenthood at birth. Similarly, the higher rates of single parenthood among low-SES household reflect the higher separation rates throughout childhood, rather than a difference in the likelihood of single-parenting at birth. Separation rates are monotonically decreasing with SES at every single age during childhood. Interestingly, whereas separation rates are falling with age among low-SES households, separation rates are increasing with age among children in high-SES households.

Third, we show that the distinction between divorce and parental separation is critical for establishing the descriptive facts. Fewer than 40% of parental separation events among recent birth cohorts involve a divorce, which means that the median parental separation event does not actually involve a divorce. Since the selection of low-SES households into marriage has become increasingly positive over recent decades, relying on divorce underestimates the SES-gradient in parental separations by 50%. Furthermore, the monotonically increasing SES-gradient in parental separations over the last decades is virtually invisible when using divorce. As such, divorce has become a poor indicator to study developments in parental separation events in recent decades.

Fourth, we show large heterogeneity in parental separation events across space. Whereas parental separations were concentrated in large cities and the suburbs between 1950 - 1970, the gravity of parental separations events has shifted towards smaller, formerly industrial towns. In the most affected parts of Sweden, over half of children born in recent cohorts experience a separation event during childhood. Differences in parental characteristics can explain less than 30% of the spatial variation, highlighting that the suggesting that the local environment and community may play an important role for prevalence of parental separation events.

These patterns raise the question why family stability has diverged so strongly across socioeconomic groups. Using the variation across space and time, we link changes in household stability to changes in household formation, economic conditions, and the outside options of both males and females. The results suggest that a large share of the changes in household stability can be explained by the decline in labor market prospects of low-educated males, in particular in more rural and remote locations. Changes in assortative matching and the improvements in female labor market prospects appear to be less important in explaining the changes.

Beyond the differences in separation probabilities at the group level, we also study heterogeneity in the parental separation probability at the individual level. To predict separation events, we train a generalized random forest model (Athey, Tibshi-

rani and Wager, 2019) based on the universe of at-birth characteristics observed in the administrative data for each parent and the match. The model performs well out of sample and predicts parental separation probabilities ranging from 20% to 75%, highlighting that parental separation probabilities vary widely between households. At the same time, even children born into the most economically and demographically favorable families face a high baseline separation risk that appears difficult to avoid based on economic and demographic characteristics.

Finally, we study the effects of parental separation events on the educational outcomes of children. We use the variation in the timing of parental separation events between siblings to estimate the effects. We find limited effects of parental separations on the national test scores in grades 3, 6, and 9, despite these being highly predictive of future educational outcomes. Parental separation events have the strongest effect when they occur shortly before the test, which suggests disruption effects, rather than a persistent effect on learning outcomes. These estimates are in line with cross-sectional evidence, with children whose parents separated prior to the various national tests performing 0.03 - 0.07 SD worse compared to non-separated parents, conditional on SES. In comparison, a one year increase in maternal education is associated with a 0.10 - 0.17 SD increase in test scores. As such, parental separation events appear to have limited effects on the learning outcomes captured by the national test scores.

The impact of parental separations on educational outcomes significantly changes when focusing on teacher-assigned grades. Children whose parents separated at any point prior to age 15 score 0.10 - 0.15 SD worse in terms of compulsory school GPA compared to their siblings who completed compulsory school at the time of the separation event. The cross-sectional difference in teacher-assigned grades similarly is three times larger than the difference in national test scores, despite the fact that the compulsory school GPA and grade 9 national test scores are available for the same children, for the same subjects, and are determined very close in time.

Given that compulsory school GPA is the sole admission criterion for secondary schools, the negative effects on teacher-assigned GPA affects subsequent educa-

tional outcomes. Experiencing a parental separation event prior age 16 lowers academic track enrollment by 2 - 3 pp compared to their siblings who were somewhat older at the time of the separation. These estimates are to the findings by Hertegård (2025*a*) based on a Swedish divorce law reform in 1974. The decline in academic track enrollment is fully explained by the lower primary school GPA, with little evidence that parental separations affect track choices conditional on GPA. As such, it is the effect on teacher-assigned grades, rather than an effect on learning outcomes as captured by standardized tests or changes in educational choices that explain the lower levels of secondary school attainment among children whose parents separated during childhood.

Our paper contributes to three strands of literature. First of all, we contribute to the literature that seeks to understand the prevalence and impact of parental separations. We extend the literature by using detailed administrative data including residential location over a 70-year period, which enables us to move beyond the studying parental divorces as is typical in the literature (Persson, 2020; Frimmel, Halla and Winter-Ebmer, 2024; Hertegård, 2025*a*; Brown, Flinn and Mullins, 2025). As we show, using parental divorce as a proxy for all types of parental separations can lead to misleading conclusions regarding the incidence and developments in parental separations. We also contribute to the literature on the impact of parental separations (Gruber, 2004; Hertegård, 2025*b*; Johnston, Jones and Pope, 2025) by providing evidence on how parental separations shapes the human capital outcomes of children. Our results highlight that parental separation events have little impact on learning outcomes as measured by test scores, but much larger effects on teacher-assigned grades. This suggests that school-based interventions, rather than household interventions, may be most successful in mediating the effects of separations on children's educational outcomes.

Second, we contribute to research on the determinants of inequality of opportunity. Previous studies show the impact of a variety of factors in shaping equality of opportunity, such as differences in place of residence (Chetty and Hendren, 2018; Chyn, Collinson and Sandler, 2025), local school financing (Biasi, Lafortune and Schönholzer, 2025), and early educational opportunities (Cornelissen et al.,

2018; Gray-Lobe, Pathak and Walters, 2023). Recent studies also investigate the role of negative childhood shocks that are typically more common among low-SES households, such as evictions (Collinson et al., 2025) and out-of-home placements (Helénsdotter, 2025). In this paper, we highlight the role of a different shock: parental separations. The results suggest that changes in household stability over recent decades are an important factor to explain the enduring educational disadvantage experienced by children from low-SES backgrounds as recently highlighted by Chetty et al. (2026).

Third, we contribute to a recent literature that attempts to understand more generally the changes in family formation and separation events over time (Reynoso, 2024; Dias et al., 2025) by studying the determinants of parental separations at the group level and individual level. Using rich administrative data, we shed light on the factors that explain the deterioration in the stability of low-SES families over time, while leaving the stability of high-SES families virtually unaffected. Our findings appear to support the notion of Autor, Dorn and Hanson (2019) in the European context, namely that changes in economic conditions among low-SES males appear to be an important driver in explaining the changes in family stability over recent decades. Finally, beyond highlighting the individual level predictors of separation events, our prediction model shows a high baseline separation probability even among families with the most favorable economic and demographic characteristics. This suggests that idiosyncratic relationship factors remain an important driver of parental separation events, in addition to structural factors related to demographic and economic changes.

The remainder of this paper proceeds as follows. Section 2 discusses the data and methodology. Section 3 provides descriptive evidence regarding the developments in parental separations and its relationship with SES. Section 4 presents a model of parental separations to explain the observed facts. Section 5 presents the results of the generalized random forest to highlight the predictors of separation at the individual level, whereas section 6 estimates the effects of parental separations on educational outcomes. Section 7 discusses implications and concludes.

# 2 Data and sample selection

We use administrative data provided by Statistics Sweden that covers the period between 1968 - 2023. The data covers the full Swedish population, and contains information on educational attainment, income, residential location, migration background, as well as other various characteristics. We are also able to link parents to children across multiple generations.

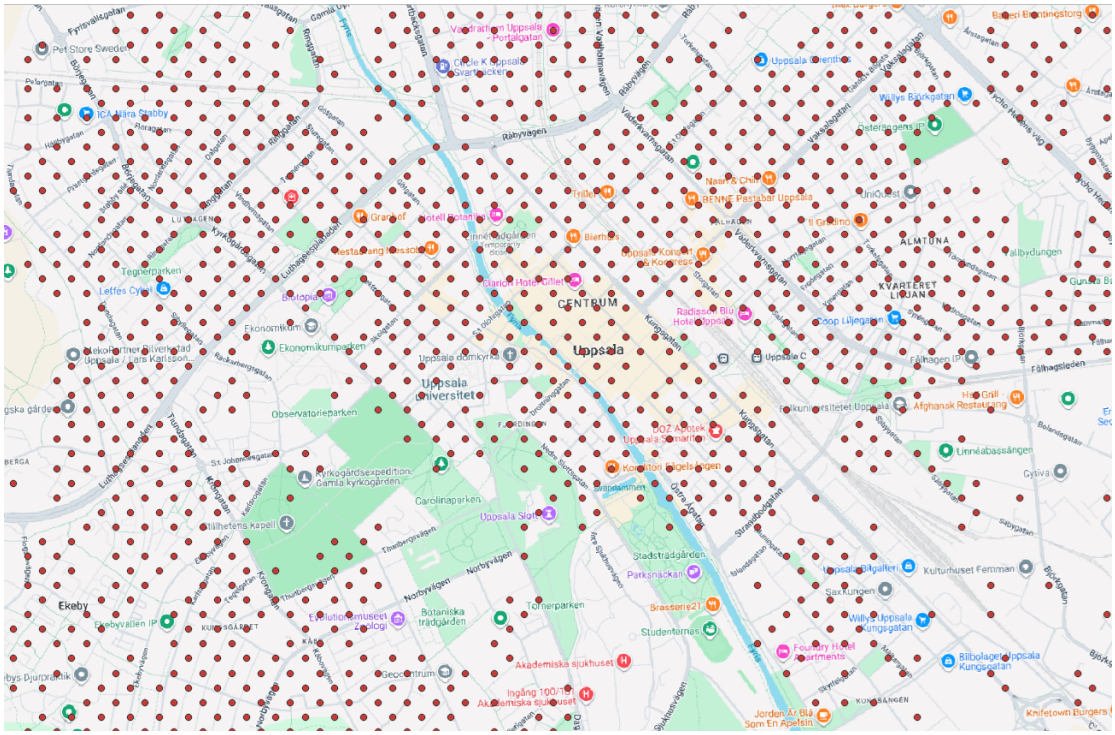
## 2.1 Parental cohabitation and separation

A novelty of our data is that it records the residential location of the Swedish population based on a grid consisting of 100 x 100 meter squares on December 31st of every year starting in 1982. Figure 1 shows the granularity of the residential location data for the city of Uppsala. We classify parents as cohabiting in a given year if they are registered in the same 100 x 100 meter square at the end of the year. Although this approach might classify separated parents as cohabiting if they live in very close proximity to each other, in practice the grid is fine enough that such cases are very few. We show this by comparing our measure of cohabitation based on the residential location to cohabitation according to the Register of the Total Population for the later cohorts in our sample. For later cohorts, the Register of Total Population classifies parents as belonging to the same family unit if they either are married or are cohabiting with common children. Both measures of cohabitation overlap almost perfectly in the years during which they are both available as shown in appendix Figure A.1.

For the earliest cohorts in our sample (those born in 1952 - 1965), we measure cohabitation based on the Register of the Total Population. In the earlier years, only married individuals are considered as being part of the same household.[1] While this limits some of our analysis for these earlier cohorts, our preferred geographic location data based on residential colocation fortunately become available just before the rise of unmarried cohabitation in Sweden. Since much of our anal-

[1]From 1998 onwards, cohabiting individuals with common children also begin to be classified as part of the same household in the Register of the Total Population. We can already observe these from 1982 onward using the location data.

Figure 1: Granularity of residential location data




*Note*: The figure shows the spatial scale on which place of residence is recorded for the city of Uppsala. Each dot represents a grid-cell of 100 x 100 meters. Grid cells without any registered residents are excluded.

ysis considers parental cohabitation at age 16, and having unmarried cohabiting parents at age 16 was still very rare for children born in 1966, the measurement error is limited, and no trend-breaks in cohabitation are visible once we switch to the residential data in 1982, as shown in appendix Figure A.1.

We classify couples as having separated if they cease cohabiting and do not cohabit again at a future point in time. We classify those who begin cohabiting again as not having separated in order to avoid conflating true separations with cases where parents temporarily reside in different locations for e.g. work or other reasons. The drawback of this approach is that true separations, where the couple separates for a few years and gets back together later, are not classified as separations in our analysis. However, cases where a couple ceases cohabiting and then begins to cohabit again for a sustained period of time are relatively rare.

## 2.2 Definition of SES

Since we are interested in studying the developments in separation events by SES, it's important to discuss our definition of SES. Throughout the baseline analyses, we consistently define SES based on the years of maternal education. The reason for this is three-fold. First of all, maternal educational characteristics are virtually always available and can be linked to children. In contrast, particularly among single-parent households at birth, the details of the father are not always known. Given that the likelihood of this censoring differs by SES, excluding all children with missing details on the father would introduce a SES-based bias. Second, the years of maternal education are highly predictive of the outcomes of children as shown in appendix Figure A.2. Among mothers with just compulsory education, 20% of children enroll in university. 80% of children born to a mother with a 4-year college degree enroll in university. As such, maternal education and children's outcomes are highly correlated. Third, as recent literature points out, maternal characteristics are substantially more predictive of the outcomes of children compared to paternal characteristics (Lundborg, Plug and Rasmussen, 2024). As such, assigning families to SES based on maternal characteristics is more closely linked to the outcomes of children, compared to paternal characteristics.

Throughout the paper, we transform maternal education into years of education. The conversion of the educational attainment to years of education is discussed in detail in appendix section B.1. Information on educational attainment of mothers becomes available from 1990. For the cohorts born in 1952 - 1980, we define maternal education based on what was observed in 1990. For later cohorts, we define maternal education based on what was observed 10 years after childbirth.

## 2.3 Sample selection

We present changes over time in parental cohabitation rates for the cohorts of children born in 1952 - 2007. We end this analysis with the 2007 cohort because this is the last cohort for which we observe children during the year in which they turn 16. This is an important time in the Swedish context, as it is when children typically graduate from primary school and complete compulsory education.

Beyond showing changes over time the longest possible time-period, we also present more detailed results for the cohorts born between 1995 and 2007. We focus on this group for two reasons. First, there are substantially more extensive register data on parental income and education starting in 1990, which means that for children born in this period, we observe at least 5 years of parental characteristics pre-birth. Second, parental separation rates were relatively stable across these cohorts, unlike the rapid decline in family stability across the preceding decades. This means that we are able to present results without worrying too much about changes within the sample period. Characterization of couples by separation probability using machine learning is done for a subset of these children (those born in 1995 - 2000). Finally, we observe detailed data on educational outcomes, including a variety of national test scores and primary school GPA, for the majority of these cohorts.

We restrict our sample to children born in Sweden to Swedish-born mothers. There are four reasons for this restriction. Firstly, we do not necessarily observe age at parental separation for foreign-born children. Second, data on parental characteristics (including educational attainment) is of higher quality and more comparable for children who have Swedish-born mothers. Third, focusing on children of Swedish-born mothers allows us to abstract from changes in ethnic and religious composition of the population, due to the substantial migration inflows to Sweden during the period under study (even within the shorter 1995 - 2007 period). Finally, we are able to make intergenerational connections to the grandparent generation for the overwhelming majority of children with Swedish-born mothers. We show robustness checks including children of foreign-born mothers for our key findings, although this has generally little impact on the results.

Finally, we exclude children who emigrate from Sweden or are deceased before the year they reach the age of 16. Our final sample for showing descriptive changes over time consists of 5,109,832 children born between 1952 and 2007. Changes in cohabitation rates by maternal education are presented for those whom maternal education is known. As such, children with mothers who had died or emigrated before 1990 have missing information on maternal education and are excluded

from the corresponding results. This excludes at most 13% of children (in the 1952 cohort). In total, the sample used for estimating these results consists of 4,937,052 children. For the more detailed analyses based on the 1995 - 2007 birth cohorts, we have complete data on 1,035,740 children.

Throughout the descriptive analyses, we consider the child to be the unit of observation. As such, the analyses and results reflect the experiences of children, rather than parents.[2] In our analysis of the predictors of parental separation, we instead consider the parental match as the unit of analysis, regardless of how many children the couple have. This analysis is therefore based on parental matches where the mother is born in Sweden and where the partners have their first child in 1995 - 2007. We use 229,894 parental matches.

# 3 Descriptive evidence on parental separations

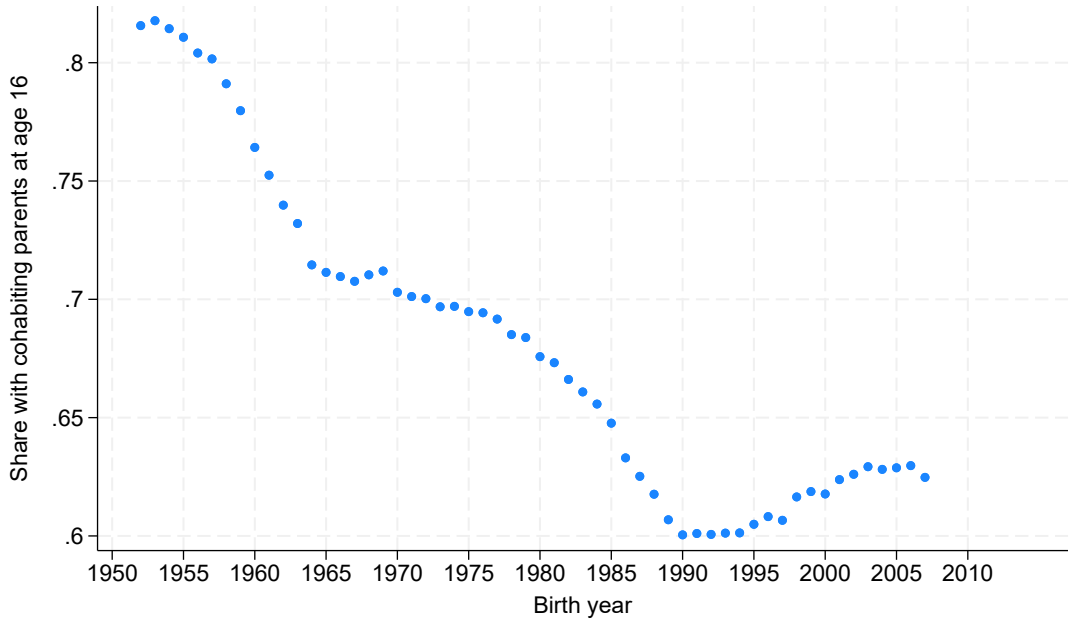
To understand the evolution and nature of parental separations, we begin by documenting several descriptive facts that guide the rest of the analysis. Specifically, we document (1) how separation rates have evolved across socioeconomic groups, (2) the type of separation events that occur, (3) the importance of the distinction between separations and divorce, and (4) the large spatial variation in separations.

## 3.1 Developments in parental separations

Figure 2 shows how the likelihood that children have cohabiting parents at age 16 has developed among cohort born between 1952 - 2007. Among children born in the early 1950s, over 80% of children had cohabiting parents at age 16. At the same time, a gradual decline in household stability is already apparent among the first cohorts for which we have data, which gathered speed between the 1957 and 1964 birth cohorts. Among those born in 1964, 71% of children had cohabiting parents at age 16. Cohabitation rates continued to decline at a slow pace until the the late 1970s. By the 1990, only 60% of children born in Sweden had cohabiting parents at age 16. Cohabitation rates remained at this low level for children born

[2]In the appendix we separately show results when restricting the sample to first-born children in case the age of the oldest child in the household might affect family dynamics. We generally find very similar estimates in the full sample and when focusing on first-born children.

Figure 2: Parental cohabitation at age 16 by birth cohort




*Note*: Figure shows parental cohabitation probability when the child is aged 16 for the birth cohorts 1952 - 2007. Cohabitation measured as parents being married according to the Register of the total population for cohorts born in 1952-1965 and as parents living in the same 100x100 meter cell for cohorts born in 1966-2007. See Figure A.1 for a comparison of sources on cohabitation rates.

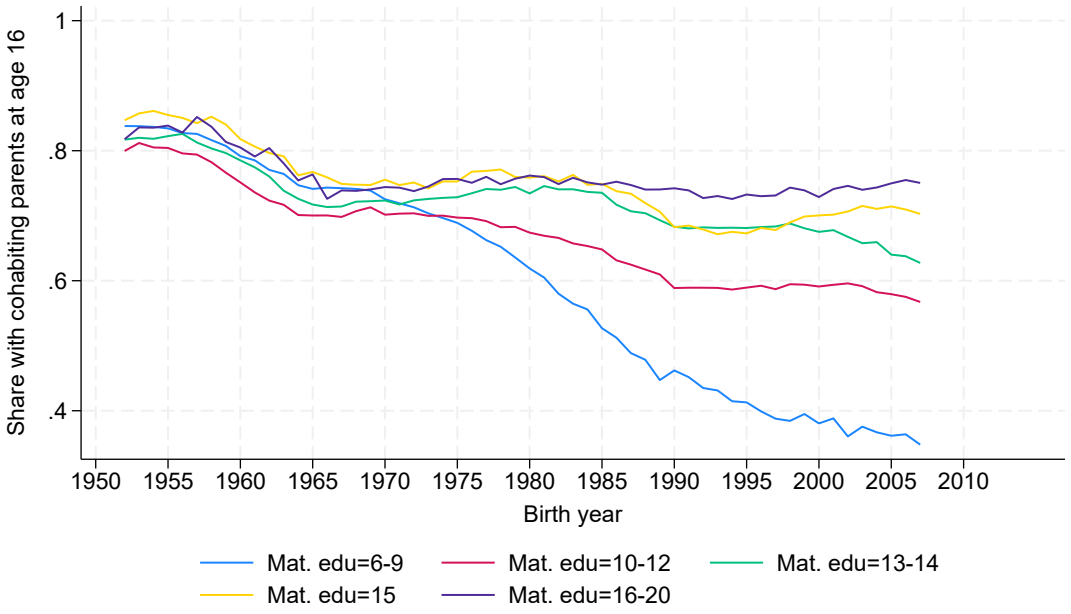
in the early 1990s, until there was a slight rebound in the late 1990s. Among children born in the most recent cohorts for whom data is available, around, 63% of children had cohabiting parents at age 16.[3]

Panel (a) of Figure 3 breaks down the developments in parental cohabitation rates by maternal education. Among the early cohorts in our sample, a similar share of children (80 - 85%) grew up with cohabiting parents regardless of maternal education. The initial decline in household stability in the late 1950s up to the early 1970s was shared between all maternal education groups.

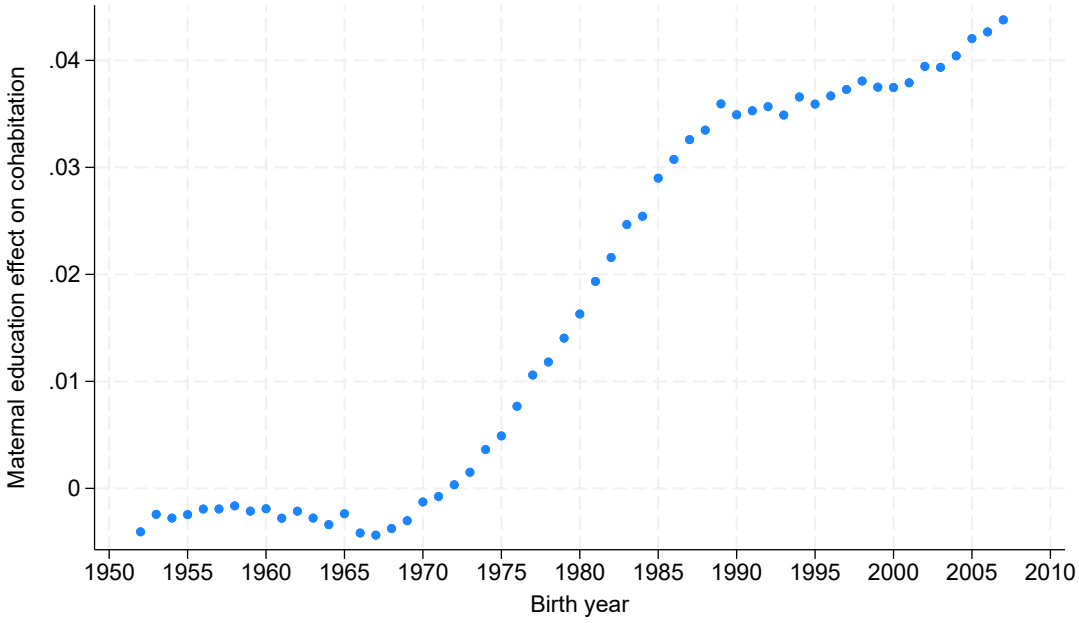
However, this changed starting with the cohorts born in the 1970s, when the stability of highly-educated households stabilized, while the stability of low-educated

[3]While these Figures may appear high within an international context, they are completely consistent with official data from Statistics Sweden (SCB), who report that 3 - 3.5% of children between ages 0 - 17 experience a parental separation event each year. See: https://www.scb.se/en/finding-statistics/statistics-by-subject-area/population-and-living-conditions/living-conditions/children-and-their-families/pong/statistical-news/children-and-their-families-in-2022

Figure 3: Changes in household stability by maternal education




(a) Parental cohabitation at age 16 by birth cohort and maternal education




(b) Slope of the SES-gradient in parental cohabitation at age 16 by birth cohort

*Note*: Panel (a) shows share of parents cohabiting when the child is aged 16 for different cohorts by maternal education. Panel (b) plots the regression coefficient obtained when regressing parental cohabitation at age 16 on years of maternal education for each birth cohort between 1952-2007.

households continued to deteriorate. For children born to a mother with 10 - 12 years of education, the likelihood of having cohabiting parents at age 16 dropped from 70% in 1970 to 58% for those born in the mid-2000's. The decline in household stability for mothers with only compulsory education has been even more severe. In contrast, the likelihood of having cohabiting parents hardly changed between 1965 and 2007 for children born to mothers with a college degree. As such, the overall decline in cohabitation rates during this period has almost entirely been driven by children born in low-SES households.

One concern is that the changes by education level reflect changing selection into educational categories since mothers became more educated over time, rather than changes in the experiences by maternal education. However, this is not the case. Appendix Figure A.4 replicates panel (a) of Figure 3 based on the relative position of the mother's education within the birth cohort. The results are very similar, highlighting that it is the relative maternal education, rather than the exact level, that drives the widening SES-gap in household stability over time.

Panel (b) of Figure 3 provides a parametric estimate of how the relationship between maternal education and parental cohabitation developed over time. Each dot is obtained by regressing the parental cohabitation rate at age 16 on maternal education for each birth cohort. The gradient with respect to maternal education was actually negative for children born before 1970. For these cohorts, one additional year of maternal education was associated with a 0.2 - 0.3 percentage points lower probability of parental cohabitation at age 16. The gradient reversed around 1970, and has monotonically increased ever since For the 2007 birth cohort, an additional year of maternal education was associated with an increase in the likelihood of parental cohabitation of more than 4 percentage points.

## 3.2 Types of parental separations

Figure 2 shows that the share of children who do not live with both parents at age 16 increased significantly over time and accounts for 37 - 40% of children born since 1990. In addition, Figure 3 shows that large differences in family stability have appeared between children with high- and low-educated mothers. To study

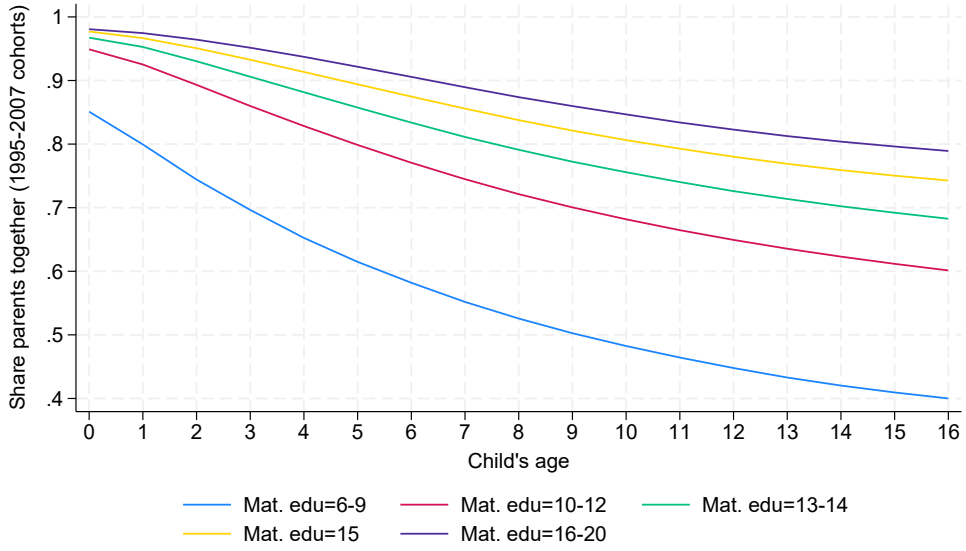
these development in more detail and to distinguish between the role of parental separation events during childhood and single parenthood at birth, Figure 7 shows the likelihood of parental cohabitation and the separation rate (i.e. hazard rate) by age and SES for the 1995 - 2007 birth cohorts.

Panel (a) shows the likelihood that parents cohabit at each age by maternal education. The differences in single-parenting at birth are comparatively small, with 95 percent of parents cohabiting at birth among the 10 - 12 years education category and 98 percent of parents cohabiting at birth among the 16 - 20 years education category. The only group which stands out are children whose mothers have at most 9 years of education. Among this group, 85% of parents cohabit at the child's birth.

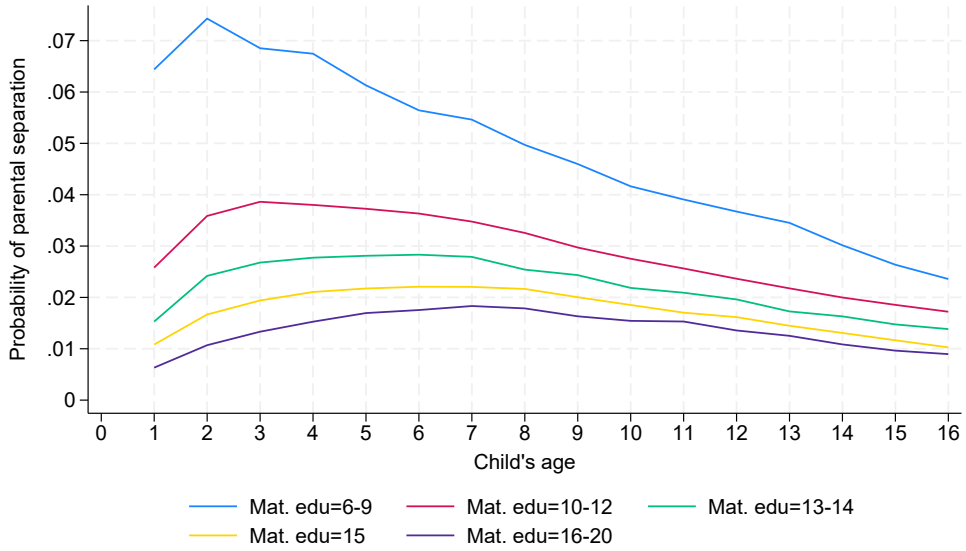
The majority of the differences in parental cohabitation at age 16 by SES are thus driven by differences in separation rates during childhood. This is confirmed by Panel (b) of Figure 4, which plots the likelihood that parents separate at any given age, given that they were still cohabiting the year before (i.e., the hazard rate). The likelihood of experiencing parental separation is higher for children of less-educated mothers at every single age in childhood. The hazard rates at every age are monotonic in terms of maternal education, despite a smaller share of low-SES parents cohabiting at each child age due to the higher separation rates experienced up to that age.

The slope of the age-gradient reveals further information. While the hazard rates are steeply downward sloping with age for children born to low-educated mothers, they are much flatter for households where the mother is highly educated. The peak in hazard rates occurs at a much later age for households where the mother is highly educated. This indicates that the mechanisms that underlie parental separation events may differ between socioeconomic groups. Among low-educated mothers, the pattern is consistent with the existence of a large share of low-initial surplus couples that break up soon after they have a child due to information updating or adverse shocks. In contrast, the patterns for high-SES households are suggestive of a process of slow deterioration following an initially high match surplus.

Figure 4: Parental colocation probability and separation hazard rate by age




(a) Parental co-location probability by age and maternal education




(b) Separation hazard rate by age and maternal education

Note: Figure shows parental cohabitation probability at ages 0 - 16 by years of maternal education (Panel a), and parental separation probability by age by years of maternal education (Panel b). Figure based on children in our sample born in 1995 - 2007.

## 3.3 Parental separation versus divorce

A key innovation in this study is the identification of parental separations based on the end of cohabitation, rather than divorce. To highlight how these two mea-

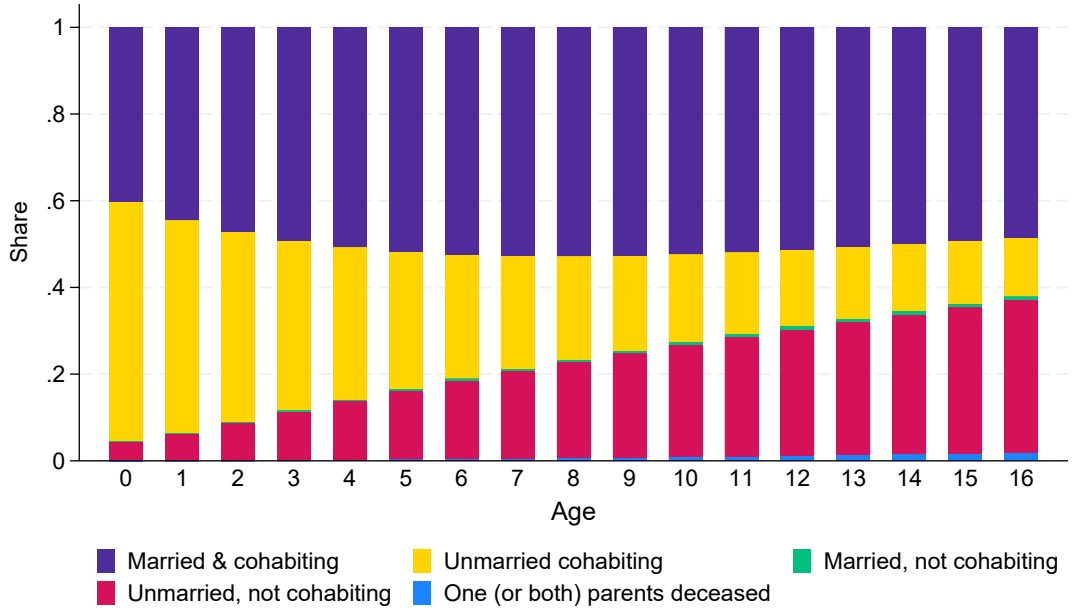
sures differ, this section shows the selection in observed separation events when exclusively focusing on divorce, which is the most commonly used measure in the literature. While unmarried cohabitation became common in Sweden earlier compares to most other countries, it is increasingly common across Europe. For instance, 58% of children in France and 50% of children in Spain are born to unmarried couples in recent years (European Commission, 2025).

Figure 5 shows the development in parental relationship status over the course of childhood for the 1995 - 2007 birth cohorts. At birth, the most common form of parental partnership is cohabitation without marriage (55%), followed by married cohabitation (40%), and unmarried non-cohabitation (4%). As such, married parents are a minority of parents at the time of birth. At age 16, the share of children with married and cohabiting parents is 48 percent, significantly higher than the share at birth. The reason is the low persistence of unmarried cohabitation as a relationship state. Among parents who are cohabiting but unmarried at birth, 37% have separated, another 37% are married and cohabiting, and only 24% remain unmarried and cohabiting by the time the child reaches age 16. As such, inflows into marriage among unmarried cohabiting couples more than offsets divorces among married couples in the earlier years of childhood.

However, focusing on divorce severely skews observed separation events towards high-SES families, since low-SES mothers are much more likely to be unmarried and cohabiting. This is shown in detail in appendix Figure A.5. While over half of parental separation events among mothers with a college degree involves a divorce, only 20% of separation events among mothers with the low levels of education involve a divorce.

Figure 6 shows how this distinction between parental separations and divorce affects the observed SES-gradient. The figure shows the maternal education gradient with respect to cohabitation, separation, and divorce at age 16. Patterns regarding cohabitation and separation are similar, since the vast majority of non-cohabitation in childhood is due to separation of initially cohabiting parents as highlighted in the previous section. In contrast, the relationship between SES and divorce is substantially weaker. This is driven by the positive selection of low-SES

Figure 5: Parental relationship structure by age




Note: Figure shows parental relationship structure by age for the birth cohorts 1995 - 2007. If parents are unmarried and not cohabiting at an earlier age, but cohabiting (either unmarried or married) at a later age, they are reclassified as unmarried cohabiting at the earlier age. If parents are married and not cohabiting at an earlier age, but are married and cohabiting at a later age, they are reclassified as married cohabiting at the earlier age.

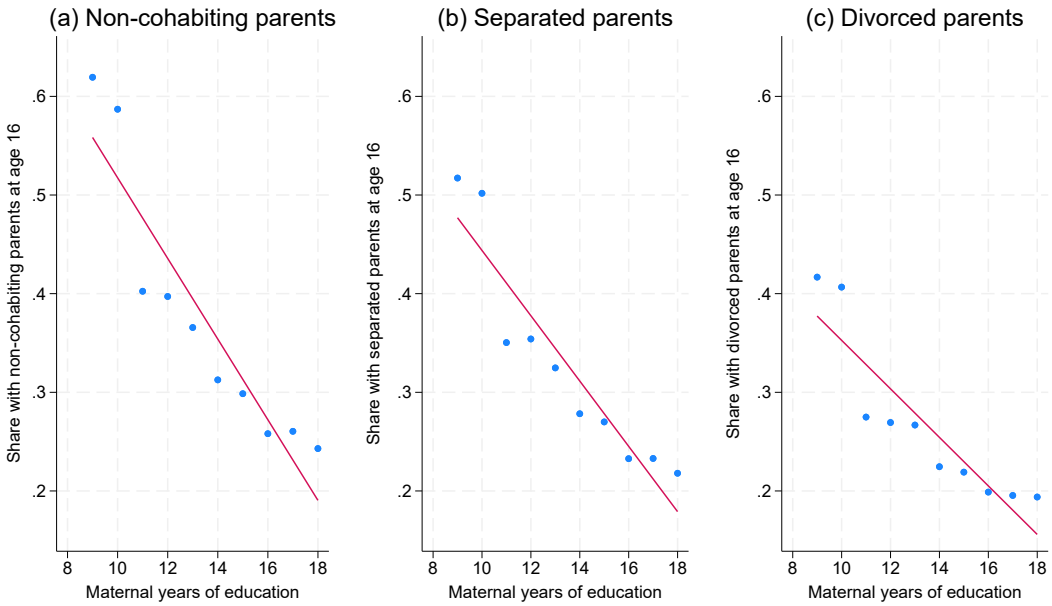
households into marriage. Whereas the parental separation probabilities for mothers with 16 years or more of education are similar among married and unmarried households, a substantial difference exists for mothers with a high school degree or less, who make up 52% of the sample.

Furthermore, Figure 7 shows that even the increasing association between maternal education and parental separation probability discussed in Figure 3 is far less prevalent when restricting attention to parental divorce. As such, considering parental separations that do not involve a divorce is critical to understanding societal changes over the last decades and to reflect parental separation rates among low-SES households.

## 3.4 The changing geography of separations

Finally, we highlight the existence of substantial spatial variation in parental separation probabilities, which suggests that local economic or social environments

Figure 6: SES gradient in parental non-cohabitation, separation, and divorce

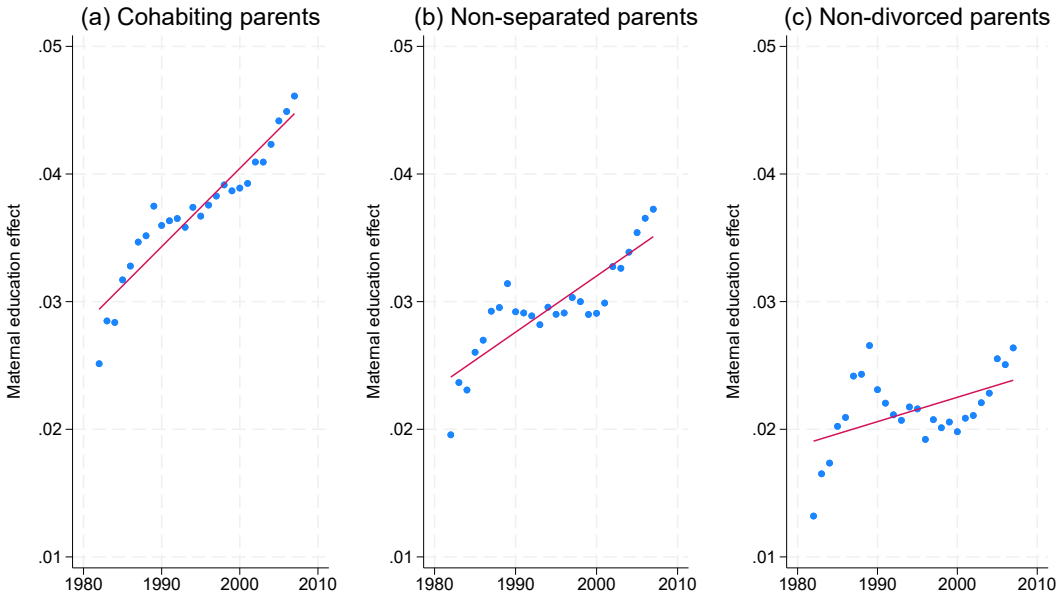



Note: Panel (a) shows the gradient between maternal education and the probability of having non-cohabiting parents at age 16. Panel (b) shows the gradient between maternal education and parental separation by age 16, conditional on the parents having cohabited at birth. Panel (c) shows the gradient between maternal education and having experienced divorce by age 16, conditional on parents having been married at any point prior to age 16. The respective slopes are -0.0402 (non-cohabitation), -0.0315 (separation) and -0.0218 (divorce). Figure based on children in born between 1995 - 2007.

may play an important role in shaping household stability. Figure 8 shows the probability that parents cohabit at age 16 by population density for the 1970 - 1972 and 2005 - 2007 birth cohorts. Despite the absence of an SES-gradient in separations for the 1970 - 1972 cohort (see Figure 3), substantial spatial variation in separation probabilities exists. Over 80% of children born in 1970 - 1972 to a household in rural communities had cohabiting parents at age 16. In contrast, the parental cohabitation rates at age 16 were well below 70% in denser places. As such, even though the variation by SES is limited for the 1970 - 1972 birth cohort, we see a clear spatial gradient on the urban-rural dimension in household stability.

The geographical patterns are substantially different 35 years later. Among the 2005 - 2007 birth cohorts, children growing up in rural communities are now significantly less likely to live with cohabiting parents at age 16 compared to children

Figure 7: Changes in SES gradient in parental cohabitation, non-separation, and non-divorce over time



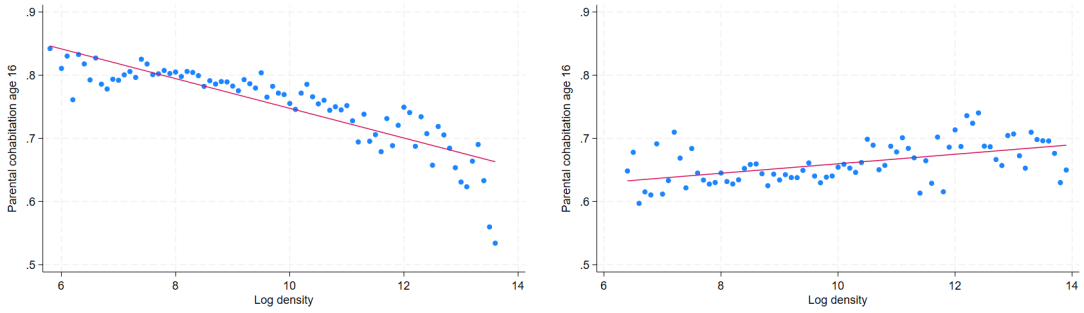

Note: Panel (a) shows the gradient between maternal education and the probability of having cohabiting parents at age 16 by birth cohort. Panel (b) shows the gradient between maternal education and the probability of not experiencing a separation of previously cohabiting parents by age 16 by birth cohort. Panel (c) shows the gradient between maternal education and the probability of not experiencing a divorce, among parents that were married at any point prior to age 16 by birth cohort. Figure based on children born between 1982 - 2007, since we can only identify single-parents at birth based on geo-coded location data from 1982 onward.

in urban communities. Interestingly, there has been virtually no change in household stability for places at the top end of the density distribution, with around 70% of children cohabiting in denser areas. As such, virtually the entire change in parental separation rates during this period is driven by a decline in household stability in rural areas, while leaving the separation rate in cities virtually untouched.[4]

The flatter urban-rural gradient for the recent birth cohorts does not imply that spatial variation no longer exists. Figure 9 shows a map with the likelihood that parents cohabit at age 16 for each municipality in Sweden. Among recent birth co-

[4]This is not simply a function of suburbanization or white flight during this period. The suburbs of Stockholm fill virtually all other top-10 spots for the 1970 - 1972 cohorts that are not occupied by the three largest cities, while none of the suburbs even break into the top-40 municipalities by separation probability for the 2005 - 2007 birth cohorts.

Figure 8: Parental cohabitation at age 16 by population density across time




(a) 1970 - 1972 birth cohorts

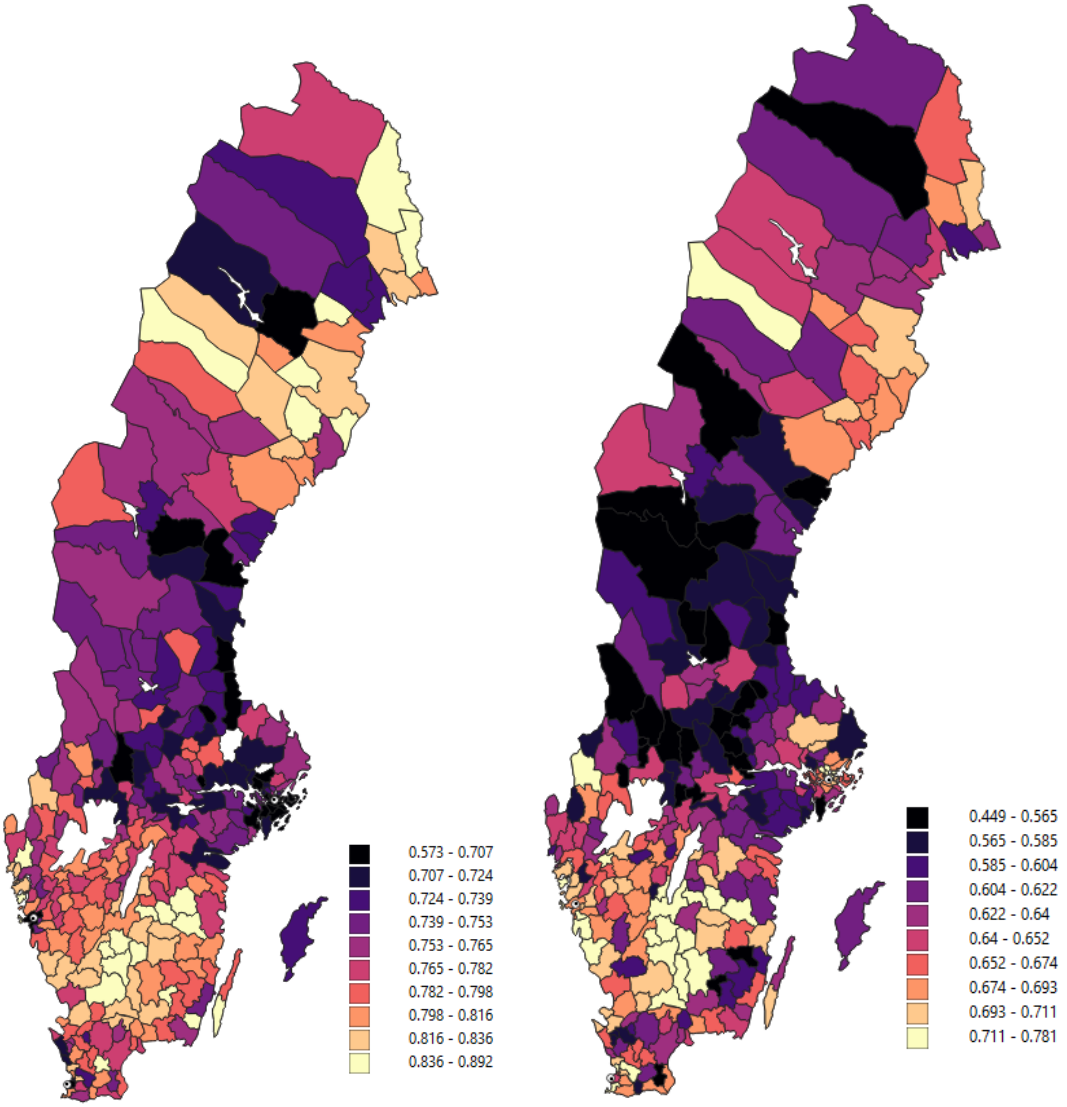
(b) 2005 - 2007 birth cohorts

Note: Figure shows how the probability of parental cohabitation at age 16 by population density for children born in the 1970 - 1972 cohort (left) and children born in the 2005 - 2007 cohort (right). Density is calculated as the log of the number of people living within 10km of the 100 x 100 meter grid-cell of residence at age 15. As such, the density measure is highly granular and also captures the substantial variation in density within municipalities. Estimates shown based on bins that are 0.1 logpoints wide. Bins with fewer than 200 observations are not displayed here.

horts, parental cohabitation probabilities at age 16 at the municipality level range from 45% to 78%. Most of the variation among recent birth cohorts are within relatively rural areas. Among the various municipalities in central-Southern Sweden, parental cohabitation rates are well over 70%, which is equivalent to every single mother holding a college bachelor's degree. In contrast, the center of parental separations shifted towards former industrial areas in central Sweden, with parental cohabitation rates that are below 50% in some municipalities. In those cases, the parental separation probabilities are equivalent to the entire population having at most a 2-year vocational high-school degree.

While SES is highly predictive of the parental separation probability (see Figure 3), it can explain only a limited degree of the spatial heterogeneity in parental separations. Appendix Figure A.6 plots the separation rates at the municipality level against the average residuals from a regression when flexibly controlling for parental SES and maternal age at birth. Accounting for detailed measures of SES and maternal age at birth can explain at most 30% of the spatial variation, highlighting that the spatial differences are not simply driven by spatial selection based on observed parental characteristics. For the 1970 birth cohort, spatial selection

Figure 9: Parental cohabitation rates at age 16 across space and time




(a) 1970 - 1972 birth cohorts (b) 2005 - 2007 birth cohorts

Note: Figure shows parental co-habitation probability at age 16 by municipality of birth for the 1970 - 1972 birth cohorts (left) and 2005 - 2007 birth cohorts (right) . To facilitate comparison, the municipalities are divided into equally-sized deciles according to the rate of parental separation by age 16. Some municipalities in Sweden are extremely small (around 2.500 inhabitants as of 2023). In both maps, municipalities with fewer than 100 children born across the three birth cohorts are excluded.

based on SES explains virtually none of the spatial variation, since SES itself was not predictive of separations (see Figure 3). As such, the spatial variation, and changing nature of the spatial variation, appears to reflect factors that extend beyond selection based on observed household characteristics.

## 3.5 Robustness of the descriptive facts

One concern is that the results in this section are based only on a subsection of the Swedish population, namely those born to Swedish-born mothers. However, including children with a migration background does not materially alter any of the descriptive facts, as highlighted in appendix section A.2. A second concern is that some of the results, in particular regarding the timing of separation events, may be driven by the presence of older siblings. In addition, the number of children to a couple is endogenous. Since each child is counted as a unique observation, this may weight the analyses in various ways. However, the descriptive facts remain unchanged if we restrict the sample to the first child born to the mother as highlighted in appendix section A.3.

# 4 Conceptual framework

### Conceptual framework

In line with the seminal work by Becker (1973) and Becker (1974), we consider that cohabiting couples form when there is a match surplus to living together compared to the outside option of each party. Beyond the direct utility generated by the partnership itself, cohabitation may also allow for intra-household specialization, meaning that the couple can produce household outputs more efficiently. While female labour market participation in Sweden has been high for most of the period under considerations, gender-typical specialization in market and non-market work nevertheless remains. For instance,Swedish women continue to take a much larger share of parental leave than Swedish men and are more likely to be employed in jobs where working hours are more flexible, e.g. in the public sector or in remote work-friendly workplaces. (Hotz, Johansson and Karimi, 2018; Ginja, Jans and Karimi, 2020). Living together also allows for the couple to specialize in different kinds of non-market household work, further expanding the household production frontier.

Couples break up when the match surplus falls below a threshold such that they obtain more utility from living separately. There are several models that can ex-

plain why match surplus may deteriorate. Becker (1974) posits that the couple might learn about their true match surplus over time, which may turn out to be less than both the expected match surplus and the outside option. Such a learning model would be observed as the probability of separation falling with the time that the relationship has lasted.

A second interpretation is that match surplus is subject to shocks which can reduce it. Such shocks could be economic in nature, such as a partner losing or finding a job, or "love shocks" that change the emotional value of the relationship. There is empirical evidence for both kinds of shocks leading to separations (Jensen and Smith, 1990; Eliason, 2012; Hess, 2004; Frimmel, Halla and Winter-Ebmer, 2024). These shocks can take two forms. One common assumption is that there is a constant probability that a relationship ends in any given year. A different assumption is that households are subject to contious shocks, leading to a random walk in the surplus from an initial baseline utility. In this case, the expected duration of a relationship strongly depend on the initial surplus.

The effects of parental separation events may depend on which mechanism underlies the separation. Partners in couples with low match surplus will invest less in family public goods than partners in couples with high match surplus due to lower expected duration of the relationship. Thus, if separations happen because partners gradually learn that their match surplus is lower than initially thought, or because an initially high relationship surplus deteriorates over time, investments in family public goods are initially high and decrease gradually. Once the separation event materializes it should have a smaller causal effect on children, as parental investments have already fallen to a lower level. Children's outcomes would already have deteriorated somewhat prior to the separation, due to the decrease in parental investments. Older children would also experience higher parental investments in their early years compared to younger children, meaning that birth order should matter more for children's outcomes in such families.

On the other hand, if separations are explained by unforeseen economic or emotional shocks, the effects on children should be more heterogeneous. Each period, some couples with high and low match surplus will break up. Effects on children

from low-surplus couples will be relatively low, while effects on children from high-surplus couples will be relatively high. There will be no deterioration in children's outcomes in periods leading up to separation, as match surplus has not systematically decreased across these periods. Furthermore, short-term disruption effects of parental separation might be present regardless of which model can best explain why separation takes place.

**Mechanisms**

Given the conceptual framework commonly used in the literature, the question is not necessarily why low-skilled households have higher separation rates, which could be explained simply in terms of lower match surplus or a higher rate of shocks arriving. Instead, the question is why the parental separation rates among low-skilled have increased so substantially throughout the last 50 years. I.e., what changes in the matching market, matching surplus, or economic shocks specifically for low-skilled could have reasonably driven the decrease in household surplus, or alternatively, changed the exit costs from a low-surplus relationship.

One possibility is that assortative matching during partnership formation has changed. If assortative matching increased and initial relationship surplus depends on the partner's SES, then this may drive the decrease in household stability among low-SES households. However, as shown in appendix Figure C.12, the degree of assortative matching did not materially change during this period. Mother from low-SES households have seen by far the largest increases in years of education among partners. When expressed in ranks, there is little evidence of a structural divergence in marriage partners quality between females of low- and high-SES over time. This is in line with evidence from the US, which shows that while assortative matching has increased over time, the magnitude of the effect remains relatively limited (Eika, Mogstad and Zafar, 2019). This suggests that changes in assortative matching are unlikely to explain the rising separation rates among low-SES households.

A second explanation is a change in social norms. While church-membership, baptisms, and church-weddings have declined across all population groups in Sweden

over the recent decades, the social norm against separations would be most binding for households with a small or marginal marriage surplus by imposing substantial exit costs. This could prevent zero or negative surplus households from separating. This mechanism lines up well with the spatial patterns, with low church membership rates in the historical manufacturing centers in central Sweden and highest church-membership rates in Southern Sweden. Viewed in this light, the rise is separation rates may not be due to a changes in relationship surplus among low-SES households, but rather, a change in the perceived social and/or reputation costs of exiting a low-surplus relationship. This result would be consistent with a higher demand for separations that has always existed among low-skilled households, but which was realized in recent decades due to the change in societal norms and attitudes. We are awaiting data regarding baptisms, church marriages, and church funerals from the church of Sweden to further explore this.

The third explanation concerns changes in the relative labor market prospects. The decline in family stability both across SES groups and space closely track the changes in the relative labor market conditions of low-SES males in the former industrial and manufacturing regions in central Sweden. Similar to the findings of Acemoglu and Restrepo (2022) in the US, automation and technological change in Sweden has predominantly affected the labor market prospects of low-SES households (Graetz, 2020). The de-industrialization process has particularly affects areas in central Sweden that historically centered around the iron industry, steel industry, and heavy manufacturing industry, which are the same regions in which the household stability among low-SES households has most severely declined. This is in line with evidence from the US explicitly linking from the decline in the local manufacturing due to import competition to a decline in marriage market value of young men, and subsequently, an increase in single-parent households (Autor, Dorn and Hanson, 2019). We are still finalizing the presentation of these analyses, and we have not yet included these results in this version of the draft.

# 5 Predicting individual separations

The previous section discusses the factors that can explain parental separation events at the group level. To gain insight into the drivers of parental separation events at the individual level, we predict the separation probability for parental matches based on characteristics of each parent and of the match. The purpose of this exercise is not causal attribution, but to characterize the extent to which parental separations are predictable from observable characteristics at birth, and to quantify residual risk even among socioeconomically advantaged households. We restrict the sample to parental matches where the first child was born in 1995 - 2007, which provides us with at least 5 years of detailed pre-childbirth data on both parents from the administrative registers. Furthermore, we restrict the analysis to parents who cohabited when their first child was born, since there is little to 'predict' based on variation in parental characteristics at birth for parents that have already separated at birth.[5]

We estimate a generalized random forests (Athey, Tibshirani and Wager, 2019) of 500 trees to predict the probability of separation based on 128 covariates observed in the administrative data.[6] The covariates include duration of cohabitation and marriage before first child, parental ages, children from previous relationships, parental education, municipality and neighborhood characteristics, as well as incomes, sectors of employment, and reliance on social payments in the five years before childbirth. Unlike a traditional prediction model, the forest bases its predictions on the variables which are the strongest predictors of divorce probability. The method is non-parametric, splitting the sample into leaves based on variables and threshold levels which maximize differences in predicted divorce probability. Our model is trained on a training set containing 80 percent of the matches in our sample, and its predictions are validated on a held-out test set containing the remaining 20 percent of matches.

The regression forest model is able to predict parental separations well in both the training set and the held-out test set as shown in appendix Figure E.15. Based

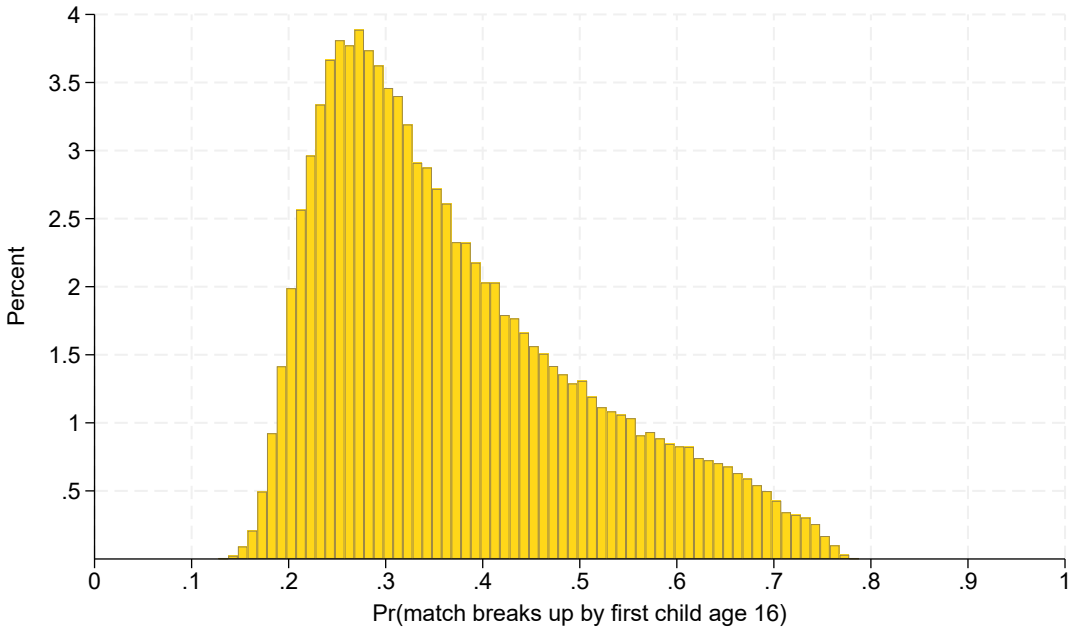
[5]With this restriction we also exclude matches where either the father or mother is unknown.

[6]The number of trees will be significantly increased in future versions of this paper.

on the held-out test set, the predicted probability of parental separation varies between less than 20 percent and over 75 percent as shown in Figure 10. Both the lower- and upper-bounds are quite interesting. At the upper-bound, the model can identify households for whom the separation risk exceeds 70%, highlighting that strong structural factors that can be identified to predict parental separation events. At that same time, very few children are identified as having a parental separation risk below 20% based on the highly-detailed household characteristics. As such, while there is a clear SES-gradient in parental separations, it appears difficult for families to isolate themselves from this risk purely on the basis of demographic and economic conditions.

Appendix Figure E.16 shows one potential measure of the importance of the

Figure 10: Histogram of predicted separation probabilities




Note: Predictions based on regression forest. Bin width=0.01.

different covariates in driving the model's predictions. The bars represent the share of times a particular variable was used by the trees of the regression forest to make splits in the data.[7] Socioeconomic variables, especially reliance on social

[7]Only splits at depths 1-4, i.e. the first sets of splits made by each tree, are included in the importance measure. Splits at a lower level receive twice as much weight as splits at the next level.

payments, education and labour income are the key variables that regression forest uses to predict parental separation probabilities.

A different, and our preferred way, to assess the importance of the variables is to identify the common set of characteristics among the households that the model identifies as having either a high or low probability of separation. Figure 11 highlights some key characteristics of parental matches with different probabilities of separating by age 16, conditional on having cohabited at birth.

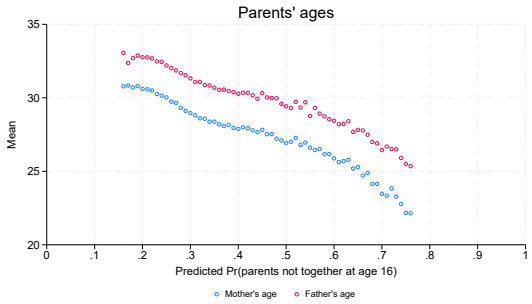
First of all, matches with higher separation probabilities are younger compared to households with lower separation probabilities. Although childbirth early in life could be interpreted as having found a high-quality match early in life within the scope of a standard search and match model, in practice, young parents are substantially more likely to experience parental separation in practice.

Second, the model highlights the predictive power of parental relationship status at birth. Among the groups with the highest separation probabilities, almost none of the parents are legally married prior to having their first child. In contrast, the average household in the low separation bins has been married for at least 2 years. Matches with a high separation probability are also disproportionately likely to have children from previous relationships.

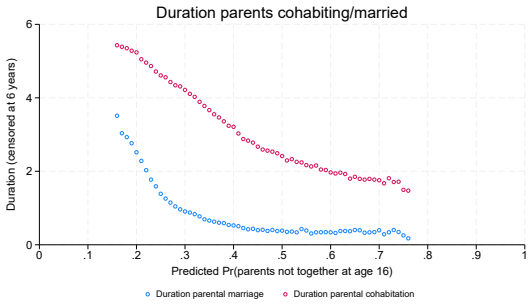
Third, those with a high separation probabilities are likely to have relatively low socioeconomic status, in line with the descriptive evidence from section 3. Parents with the highest probabilities of breaking up have typically not completed high school and have low earnings from both parents, and are more likely instead to rely on social security payments. In contrast, those with low probabilities of breaking up tend to have post-secondary education, and both parents have sizable pre-birth labor incomes.

Finally, there is substantial intergenerational persistence. Parents with low separation probabilities tend to have more educated grandparents, as shown in panel (f). Among households with the lowest separation probabilities, the average years of education of the maternal grandparents is around 10.5 - 11 years, implying at least some high-school education. In contrast, among those with a separation probability above 40 percent, the maternal grandparent on average have between 9

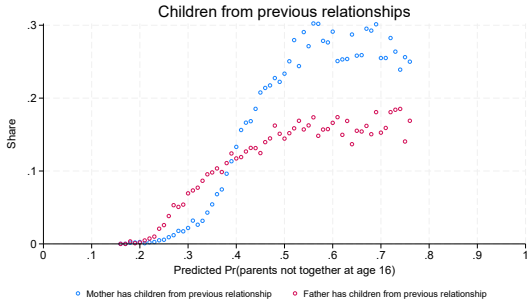
Figure 11: Descriptives of parental matches with different probabilities of breaking up



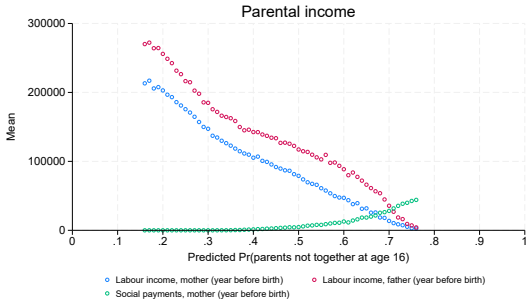

(a) Parental age at birth



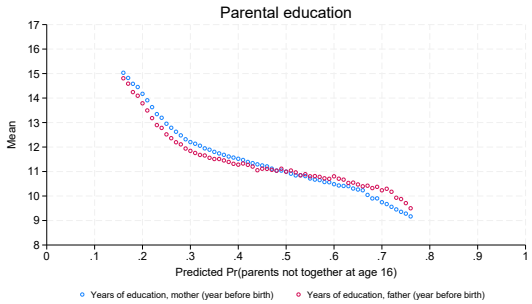

(b) Length of marriage/cohabitation



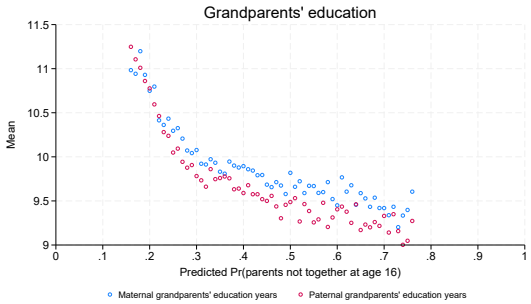

(c) Children from prior relationships




(d) Parental income




(e) Years of parental education




(f) Years of grandparents' education

Note: Means and shares of respective characteristics for parental matches in the associated percentage point of probability of breaking up, as predicted by the regression forest model.

- 9.5 years of education, which represents compulsory schooling. Interesting, these patterns are quite similar for maternal and paternal grandparent's education.

# 6 Effects of parental separations

Finally, given that we documented the trends in parental separations and highlighted the factors behind the increasing concentration of parental separations among low-SES families, we turn to estimating the effects of parental separations on the outcomes of children. Given that the separation event is endogenous to the household characteristics, both observed and unobserved, we utilize the variation in timing of parental divorce to study the impact of a parental separation event. This sheds light on the effects of parental separation on the human capital accumulation of children. Furthermore, the contrast between standardized tests and teacher-assigned grades allows us to distinguish between effects on cognitive achievement and effects mediated by discretionary evaluation within schools.

## 6.1 Methodology

Since the selection of households into parental separation differs greatly with the age at separation (see Figure 4), we rely on household fixed effect estimates to estimate the effect of parental separations that occur at various ages. In this case, we compare siblings born to the same couple who were different ages at the time of the separation event, and whom thus spent different parts of the childhood living in a cohabiting family as opposed to a separated family. We restrict attention to children born to first-time parents to avoid complications due to the potential presence of stepchildren and stepparents. Furthermore, we restrict attention to children whose parents cohabited at the start of childhood.

To estimate the effects, we estimate equation 1 for a variety of educational outcomes. The $\beta_m$ coefficients are identified of children within the same household and, who were at a different age at the time of the divorce. Equation 1 includes cohort-fixed effects, as well as controls for gender and birth-spacing. Since the $\beta_m$ coefficients are identified from within households where all children experience a separation event, we need an omitted group to compare the effects of experiencing a separation at the various ages against. We exclude children whose parents separated in early adulthood between ages 18 - 20. As such, the $\beta_m$ coefficients should be interpreted as the effect of experiencing a parental separation event at

age $m$, as opposed to experiencing a separation event in early adulthood.

$$\text{Educational outcome}_{ihc} = \alpha_h + \sum_{m=1}^{17} \beta_m \text{Separation}_{im} + \gamma_c + \Gamma X_i + \epsilon_i \quad (1)$$

One complication with the estimation of equation 1 is that there is a perfect collinearity between the difference in the birth year of siblings (i.e., birth spacing) and the difference in age at the time of the separation within households where parents separate. As such, it is impossible to independently account for birth-spacing in a sample of households where all children separate. Since the length of birth-spacing is related to the gap in cognitive attainment between siblings due to different spillovers (Price and Kalil, 2010; Buckles and Munnich, 2012), accounting for birth spacing effects may be important. To correct for the effects of birth spacing, we include children whose parents never separated in the estimation of equation 1. Since there is no variation in the age at separation within households where the parents never separate, they do not contribute to the estimation of the $\beta_m$ parameters. This assumes that the effects of birth spacing are similar within households that do and do not separate at some point during childhood. Appendix section D.1 discusses this point in more detail and provides strong support for this assumption.

## 6.2 Effect on national test scores

We start by considering the effect of parental separation on the scores on the standardized national tests during compulsory education. Sweden conducts national tests in grades 3, 6, and 9 of compulsory education to measure the educational outcomes of children. The grades 3 and 6 tests are generally considered low-stakes, since tracking does not start until after 9th grade. The grade 9 national test is somewhat higher stakes, since it informs teachers about the relative ability of students, which may feature into the primary school GPA assessments. The primary school GPA (typically assigned in the year in which children turn 16) is the key outcome for children in terms of future educational pathways, since children are allocated to their preferred high-schools and high-school tracks purely on the basis of the compulsory school GPA.

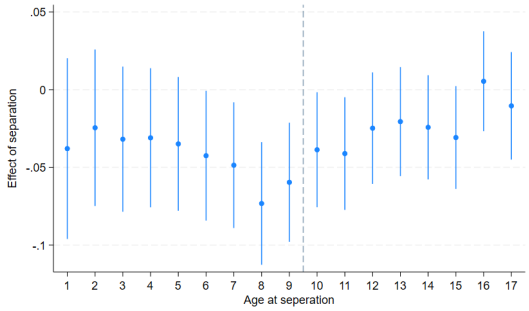
Data on the national test scores have been collected by Statistics Sweden since 2003 for the grade 9 test, 2010 for the grade 3 test, and 2012 for the grade 6 test. The grade 3 test measures children's ability in mathematics and Swedish, whereas the grade 6 and grade 9 tests also include English. Each test yields a variety of ranked sub-scores, some numerical, some categorical, and some pass/fail. We use item response theory to extract the underlying latent abilities based on the various tests in every given year. The tests generally cover a wide range of item difficulties with very few children achieving the maximum score across all sub-tests in any given year, which results in very few mass points in the inferred ability distribution. Appendix section B.2 provides more information on the distribution and predictiveness of the national primary school tests.

Panels (a) - (c) of Figure 12 shows the coefficients obtained by estimating equation 1 on the national standardized tests for all cohorts for which data is available. Three things stand out. First of all, there is limited evidence that the length of childhood spend in the pre- vs post-separation years affects learning outcomes as measured by the national test scores. Children whose parents separated during early childhood perform similar on the national tests compared to children whose parents separated much later in childhood or even in early adulthood. As such, parental separation events in early childhood appear generally not to leave long-lasting scars.

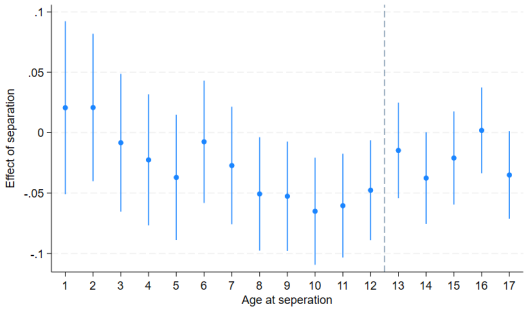
Second, there is clear evidence of disruption effects following parental separation. Across all three national tests, the effects are most negative in the few years prior to taking the test. In the case of measuring human capital outcomes for the grade 3 national test, the estimates are most negative for separations at ages 8 - 9. For the outcomes measured at grade 6, the estimates are most negative for separations occurring between ages 10 - 11. Finally, for the grade 9 national test, the effects of separation are most negative between ages 13 - 15. As such, it is not the case that the duration of childhood exposure to cohabiting or separating parents matters, but instead, the age at which the separation event happens relative to the age at which outcomes are measured.

Third, the effect sizes generally are very limited. The effect of parental separating

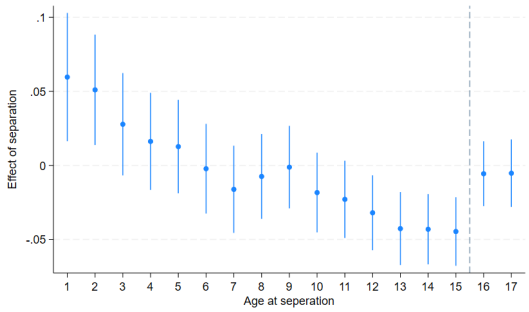
Figure 12: School outcomes and age at separation - HH FE estimates



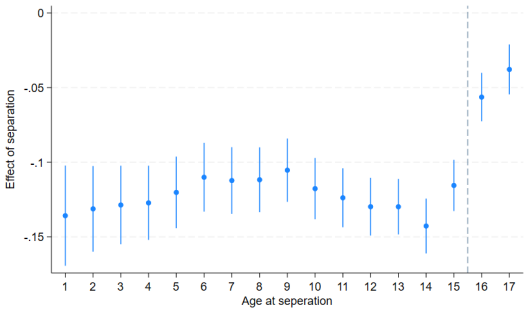

(a) Grade 3 national test



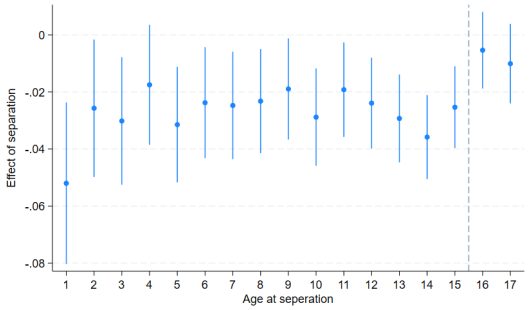

(b) Grade 6 national test




(c) Grade 9 national test




(d) Primary school grade 9 GPA




(e) Academic high-school diploma

Note: Estimated effects of age at parental separation on national standardized test scores and primary school GPA (all normalised within birth cohorts) as well as high school academic track completion. The baseline group is those whose parents separate when the child is aged 18 - 20, the estimates show the effects of experiencing a separation event at age $m$ instead of age 18 - 20. Household-fixed effects are included, so estimates are obtained from sibling comparisons. Sample restricted to children born to parents who did not have children with anyone else previously. Since the tests that place in the year that children respectively turn 10, 13, and 16, there is always a substantial group of children for whom the separation event occurs well after the outcome measurement. The dashed lines show the age at which outcomes are measured. Standard errors are cluster at the household level.

at any age in childhood on national test scores, as opposed to experiencing a separation event in early-adulthood between ages 18 - 20, is typically within the 0.05

SD range. For comparison, an additional year of maternal education is associated with a 0.11 - 0.16 SD improvement in national test scores. These effect sizes are noticeable, since a parental separation event drastically changes the household structure and household dynamics for the child from that point onward. However, it appears that the disruption caused by the change in household structure, rather than the change in household structure itself, is generating negative effects on learning outcomes.

Table 1 shows the results when we put some more structure on the individual $\beta_{im}$ estimates to compare the estimates between the different separation ages. The estimates in columns (1) - (3) confirm that it's predominantly the effects of disruption, rather than the duration of experiencing separated parents, that drives the effects on the national test scores.

Interesting, the relatively small effects of parental separations obtained from the sibling design are fully in line with cross-sectional evidence. Table 2 shows the cross-sectional differences in national test scores depending on whether parents cohabit in the year prior to the national test, both conditional and unconditional on household characteristics. Conditional on parental education and age at birth, children whose parents cohabit in the year prior to the test score between 0.04 and 0.07 SD better on the national standardized tests compared to children whose parents were not cohabiting in the year prior to the test. As such, parental separation events appear to have limited effects on learning outcomes as measured by the national standardized tests.

## 6.3 Effect on teacher-assigned grades and academic track enrollment

The effect of parental separations on educational outcomes changes substantially when focusing on compulsory school GPA. This difference is visible both in the sibling estimates as shown in panel (d) of Figure 12 as well as the cross-sectional evidence in Table 2. Rather than showing evidence of critical age effects or length of separation effects, the main distinction for the effect on compulsory GPA is whether the separation event occurred before or after the GPA was determined at

Table 1: Effects of parental separation on educational outcomes

| | (1) Grade 3 test | (2) Grade 6 test | (3) Grade 9 test | (4) Grade 9 GPA | (5) HS Track |
|---|---|---|---|---|---|
| Separated age 1 - 4 | -0.0343 | -0.0175 | 0.0117 | -0.120*** | -0.0215* |
| | (0.0207) | (0.0249) | (0.0150) | (0.0144) | (0.00959) |
| Separated age 5 - 8 | -0.0506** | -0.0369 | -0.0174 | -0.106*** | -0.0229** |
| | (0.0183) | (0.0218) | (0.0130) | (0.0125) | (0.00827) |
| Separated age 9 - 12 | -0.0417* | -0.0607** | -0.0254* | -0.117*** | -0.0213** |
| | (0.0162) | (0.0189) | (0.0112) | (0.0107) | (0.00706) |
| Separated age 13 - 15 | -0.0254 | -0.0263 | -0.0465*** | -0.130*** | -0.0295*** |
| | (0.0145) | (0.0162) | (0.00963) | (0.00916) | (0.00598) |
| Separated age 16 - 17 | -0.00257 | -0.0166 | -0.00688 | -0.0397*** | -0.00679 |
| | (0.0139) | (0.0148) | (0.00902) | (0.00835) | (0.00555) |
| $N$ | 477401 | 297076 | 595563 | 615336 | 542740 |
| $R^2$ | 0.647 | 0.716 | 0.726 | 0.738 | 0.666 |

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Estimated effects obtained by estimation equation 4 on national standardized test scores, standardized primary school GPA, and high school academic track completion. The baseline group is those whose parents separate when the child is aged 18 - 20. Household-fixed effects are included, so estimates are obtained from sibling comparisons. Sample restricted to first children born to both parents. Standard errors are cluster at the household level. See appendix section D.2 for a more detailed description and discussion of the estimating equation.

the end of compulsory education. The estimates are substantial, since experiencing a separation event at any age between ages 1 - 15, as opposed to between ages 18 - 20, is associated with a 0.1 - 0.15 SD decline in primary school GPA. This estimate is around 3 times larger compared to the effect on national test scores in both in the sibling comparison as well as in the cross-section. This is despite the fact that the compulsory school GPA and grade 9 test score are available for the same cohorts, and are determined within a few months of time.

The compulsory school GPA is the sole admission criterion on which municipalities assign children to high school programs. As such, the effect on the primary school GPA is the outcome that affects the option set available to children, rather than

Table 2: Cross-sectional difference in test scores and primary school GPA (in s.d.)

| | Grade 3 test | | Grade 6 test | | Grade 9 test | | Grade 9 GPA | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Parents cohabit | 0.155*** | 0.0733*** | 0.208*** | 0.0379*** | 0.231** | 0.0766** | 0.383*** | 0.238*** |
| Year prior to test? | (0.003) | (0.003) | (0.0021) | (0.0020) | (0.003) | (0.003) | (0.0017) | (0.0016) |
| N | 751,507 | 751,507 | 552,452 | 552,452 | 945.505 | 945.505 | 962,870 | 962,870 |
| SES controls | No | Yes | No | Yes | No | Yes | No | Yes |

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Estimates for all cohorts for which the relevant national tests scores are available. SES controls include a flexible interaction of parental education and parental age at birth. Children who were born to single-parent households are excluded from the sample. To facilitate comparison, the cohorts are restricted in case of the primary school GPA to the same for which the grade 9 standardized tests are available. Both the test scores and GPA are standardized to mean zero and s.d. 1 within each birth cohort.

the score on the national test. Panel (e) shows the effect of parental separation at the various ages on high school academic track completion, with around 60% of children in Sweden during this period attending the academic track. Children whose parents separate between ages 1 - 15 have a 2 - 3 pp lower probability to complete high-school compared to children whose parents separate between ages 18 - 20. This estimate is similar to previous estimates by Hertegård (2025*a*) based on studying the impact of parental separations using the 1974 divorce-law reform. Similar to the effect on primary school GPA, neither the duration nor age at parental separation plays much of a role, other than whether it occurs before or after age 15.

Since the compulsory GPA is the sole criterion for high school track admissions, the lower rates of academic track completion may well be due to the effect on compulsory school GPA. This is confirmed in both the sibling comparison and cross-section. As Table 3 shows, controlling for compulsory school GPA renders the estimates of the age at separation on academic track completion insignificant for all ages. Similarly, children whose parents separated prior to age 15 are around 4pp less likely to attend the academic track in the cross-section, conditional on SES. This gap fully disappears when conditioning on compulsory school GPA, highlighting that the lower compulsory school GPA, rather than different choices

Table 3: Effect of parental separations on high school academic track graduation

| | (1) | (2) |
|---|---|---|
| Separated age 1 - 4 | -0.0215* | -0.00713 |
| | (0.00959) | (0.00943) |
| Separated age 5 - 8 | -0.0229** | -0.00735 |
| | (0.00827) | (0.00816) |
| Separated age 9 - 12 | -0.0213** | -0.00304 |
| | (0.00706) | (0.00696) |
| Separated age 13 - 15 | -0.0295*** | -0.00661 |
| | (0.00598) | (0.00585) |
| Separated age 16 - 17 | -0.00679 | 0.000747 |
| | (0.00555) | (0.00537) |
| $N$ | 542740 | 488832 |
| $R^2$ | 0.666 | 0.719 |
| Conditional on Primary school GPA? | No | Yes |

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Estimated effects obtained by estimation equation 4 on high school academic track completion. The baseline group is those whose parents separate when the child is aged 18 - 20. Household-fixed effects are included, so estimates are obtained from sibling comparisons. The estimates in column (2) include a dummy for each possible score on the primary school GPA. Standard errors are cluster at the household level.

conditional on GPA, is the key driver why children whose parents separated during in childhood have lower rates of academic track completion in high school.

An important question is why the effects of parental separations differ between the primary school GPA and grade 9 national tests. The national test scores and primary school GPA are determined only a few months apart and are available for the exact same cohorts, which means that the estimates are based on the same set of students. Furthermore, we can replicate the results based on primary school GPA by focusing only on teacher-assigned grades in mathematics, Swedish, and English as shown in appendix Figure D.13. In this case, the grades are based on the same subjects that are measured as part of the grade 9 national test, which means that the differences are not driven by the subject selection. Finally, we can

replicate the difference between national test scores and teacher assigned grades based on the educational outcomes in grade 6 as shown in appendix section D.4, highlighting that it is not something particular about the high-stakes nature of grade 9 grades that drives the difference.

One possible explanation is that parental separations directly affect teacher assessments if teachers internalize the separation event as a signal of children's potential ability, in particular as parental separations are more common among worse performing students. However, a different, and more likely, possibility is that teachers grade different aspects of learning compared to the national test, including, for instance, homework preparation. Children who have separated parents may receive less effective parental support with homework and assignment preparation due to the increased coordination costs between parents in assigning homework support, keeping track of school progress, and teacher expectations. Regardless of which mechanism drives the outcomes, it appears that the effects on teacher-assigned grades, rather than direct learning outcomes, are driving the effects.

# 7 Conclusion

This paper studies the prevalence, determinants, and effects of parental separations. Parental separation is a common occurrence in childhood with over 35% of children in Sweden living in separated households by age 16, yet it remains understudied in the economic literature. In this paper, we make four novel contributions.

First, we document novel descriptive statistics surrounding parental separations. The relationship between parental separation probability and SES is a relatively recent phenomenon, with the slope of the gradient increasing from non-existent to fairly substantial over the last 50 years. This is fully driven by a decline in family stability among low-SES households. Given that the median child in low-SES households is now expected to experience a parental separation, the findings have important implications for the study of equality of opportunity.

Second, we highlight some shortcoming of the reliance of divorce as a measure of parental separations, as has been common in previous literature. The average

parental separation event does not involve a divorce for recent cohorts, given that most parents are not married at birth, and non-married parents have significantly higher separation rates. As a result, the observed separation events based on divorce in practice are strongly skewed towards high-SES households. However, since the selection into marriage changes over time by SES, the SES gradient in parental separations is underestimated by 50% when using divorce, and very little of the increasing relationship over time between SES and parental separation is visible based on divorce. As such, our approach to determine parental separations from parental colocation provides strikingly different results.

Third, we study the determinants of separations at the level of parental matches using a regression forest. The model can predict parental separations at the individual level remarkably well, with predicted likelihoods varying between 20% and 80%. The information from the model can be used in public policy to identify the most at-risk families, and consider potential interventions. Although relationships and marriages are deeply personal, previous studies show that public policies play a role in determining family stability (Gruber, 2004; Avdic and Karimi, 2018; Olafsson and Steingrimsdottir, 2020; Persson, 2020; Hertegård, 2025*a*; Adda, Pinotti and Tura, 2025). At the same time, our model also highlights some limitations. Despite the extremely rich administrative data available, children in the group with the lowest parental separation probability based on the generalized random forest still experience parental separation in about 15 - 20% of the cases. This highlights that individual characteristics and experiences play an important role, beyond the economic and social structures that can be observed in administrative records.

Finally, we provide novel evidence on the effects of parental separation on the outcomes of children. Similar to the findings of Hertegård (2025*a*) and Johnston, Jones and Pope (2025), we find effects of parental separations on the long-run educational outcomes of children. However, our data enable us to go one step further and investigate what mechanisms drive the effects on the long-run educational outcomes. We find little evidence that parental separations have permanent scarring effects on learning outcomes as measured by the standardized tests. In-

stead, we find more support for short-term disruption effects on learning outcomes following parental separations, similar to the effects of parental job loss on educational outcomes in Norway as documented in Carneiro et al. (2023). In contrast, the long-term effects are three times larger on teacher-assigned grades. Since these are critical in the Swedish context for future educational possibilities, we show that it is the effect on teacher-assigned grades, rather than effects on learning outcomes or educational choices conditional on grades, that generate the persistent effects of parental separations on long-term educational outcomes.

# References


**Acemoglu, Daron, and Pascual Restrepo.** 2022. "Tasks, automation, and the rise in US wage inequality." *Econometrica*, 90(5): 1973–2016.

**Adda, Jérôme, Paolo Pinotti, and Giulia Tura.** 2025. "There's more to marriage than love: the effect of legal status and cultural distance on intermarriages and separations." *Journal of Political Economy*, 133(4): 1276–1333.

**Andrew, Alison, Orazio P Attanasio, Raquel Bernal, Lina Cardona Sosa, Sonya Krutikova, and Marta Rubio-Codina.** 2024. "Preschool quality and child development." *Journal of Political Economy*, 132(7): 2304–2345.

**Athey, Susan, Julie Tibshirani, and Stefan Wager.** 2019. "Generalized random forests." *The Annals of Statistics*, 47(2): 1148.

**Autor, David, David Dorn, and Gordon Hanson.** 2019. "When work disappears: Manufacturing decline and the falling marriage market value of young men." *American Economic Review: Insights*, 1(2): 161–178.

**Avdic, Daniel, and Arizo Karimi.** 2018. "Modern family? Paternity leave and marital stability." *American Economic Journal: Applied Economics*, 10(4): 283–307.

**Bau, Natalie.** 2022. "Estimating an equilibrium model of horizontal competition in education." *Journal of Political Economy*, 130(7): 1717–1764.

**Becker, Gary S.** 1973. "A Theory of Marriage: Part I." *Journal of Political Economy*, 81(4): 813–1048.

**Becker, Gary S.** 1974. "A Theory of Marriage: Part II." *Journal of Political Economy*, 82(2): 1321–S233.

**Biasi, Barbara, Julien Lafortune, and David Schönholzer.** 2025. "What works and for whom? Effectiveness and efficiency of school capital investments across the US." *The Quarterly Journal of Economics*, qjaf013.

**Brown, Meta, Christopher J Flinn, and Joseph Mullins.** 2025. "Family Law Effects on Divorce, Fertility, and Child Investment." *Journal of Labor Economics*, 43(S1): S351–S397.

**Buckles, Kasey S, and Elizabeth L Munnich.** 2012. "Birth spacing and sibling outcomes." *Journal of Human Resources*, 47(3): 613–642.

**Carneiro, Pedro, Emanuela Galasso, Italo Lopez Garcia, Paula Bedregal, and Miguel Cordero.** 2024. "Impacts of a large-scale parenting program: Experimental evidence from Chile." *Journal of Political Economy*, 132(4): 1113–1161.

**Carneiro, Pedro, Kjell G Salvanes, Barton Willage, and Alexander Willén.** 2023. "Childhood shocks across ages and human capital formation."

**Chetty, Raj, and Nathaniel Hendren.** 2018. "The impacts of neighborhoods on intergenerational mobility I: Childhood exposure effects." *The quarterly journal of economics*, 133(3): 1107–1162.

**Chetty, Raj, Will Dobbie, Benjamin Goldman, Sonya R Porter, and Crystal S Yang.** 2026. "Changing opportunity: Sociological mechanisms underlying growing class gaps and shrinking race gaps in economic mobility." *The Quarterly Journal of Economics*, qjaf057.

**Chyn, Eric, Robert Collinson, and Danielle H Sandler.** 2025. "The Long-Run Effects of America's Largest Residential Racial Desegregation Program: Gautreaux." *The Quarterly Journal of Economics*, qjaf011.

**Collinson, Robert, Deniz Dutz, John Eric Humphries, Nicholas S Mader, Daniel Tannenbaum, and Winnie van Dijk.** 2025. "The Effects of Eviction on Children." National Bureau of Economic Research.

**Cornelissen, Thomas, Christian Dustmann, Anna Raute, and Uta Schönberg.** 2018. "Who benefits from universal child care? Estimating marginal returns to early child care attendance." *Journal of Political Economy*, 126(6): 2356–2409.

**Dias, Monica Costa, Pierre-André Chiappori, Costas Meghir, and Hanzhe Zhang.** 2025. "Changes in Marital Sorting: Theory and Evidence from the United States." *Journal of Political Economy*.

**Eika, Lasse, Magne Mogstad, and Basit Zafar.** 2019. "Educational assortative mating and household income inequality." *Journal of Political Economy*, 127(6): 2795–2835.

**Eliason, Marcus.** 2012. "Lost jobs, broken marriages." *Journal of Population Economics*, 25: 1365–1397.

**Frimmel, Wolfgang, Martin Halla, and Rudolf Winter-Ebmer.** 2024. "How does parental divorce affect children's long-term outcomes?" *Journal of Public Economics*, 239: 105201.

**Ginja, Rita, Jenny Jans, and Arizo Karimi.** 2020. "Parental leave benefits, household labor supply, and children's long-run outcomes." *Journal of Labor Economics*, 38(1): 261–320.

**Graetz, Georg.** 2020. "Technological change and the Swedish labor market." IFAU-Institute for Evaluation of Labour Market and Education Policy.

**Gray-Lobe, Guthrie, Parag A Pathak, and Christopher R Walters.** 2023. "The long-term effects of universal preschool in Boston." *The Quarterly Journal of Economics*, 138(1): 363–411.

**Gruber, Jonathan.** 2004. "Is making divorce easier bad for children? The long-run implications of unilateral divorce." *Journal of Labor Economics*, 22(4): 799–833.

**Helénsdotter, Ronja.** 2025. "Surviving Childhood: Effects of Removing a Child From Home." *Review of Economic Studies*, rdaf044.

**Hertegård, Edvin.** 2025*a*. *Divorce law reform, family stability, and children's long-term outcomes.*

**Hertegård, Edvin.** 2025*b*. "The Rising Educational Penalties of Parental Divorce Across Generations."

**Hess, Gregory D.** 2004. "Marriage and Consumption Insurance: What's Love Got to Do with It?" *Journal of Political Economy*, 112(2): 290–318.

**Hotz, V Joseph, Per Johansson, and Arizo Karimi.** 2018. "Parenthood, family friendly workplaces, and the gender gaps in early work careers." National Bureau of Economic Research.

**Jensen, Peter, and Nina Smith.** 1990. "Unemployment and marital dissolution." *Journal of Population Economics*, 3: 215–229.

**Johnston, Andrew C, Maggie R Jones, and Nolan G Pope.** 2025. "Divorce, Family Arrangements, and Children's Adult Outcomes." National Bureau of Economic Research.

**Kinsler, Josh, and Ronni Pavan.** 2021. "Local distortions in parental beliefs over child skill." *Journal of Political Economy*, 129(1): 81–100.

**Lundborg, Petter, Erik Plug, and Astrid Würtz Rasmussen.** 2024. "On the family origins of human capital formation: Evidence from donor children." *Review of Economic Studies*, rdae101.

**Olafsson, Arna, and Herdis Steingrimsdottir.** 2020. "How does daddy at home affect marital stability?" *The Economic Journal*, 130(629): 1471–1500.

**Persson, Petra.** 2020. "Social insurance and the marriage market." *Journal of Political Economy*, 128(1): 252–300.

**Price, Joseph, and Ariel Kalil.** 2010. "The effect of parental time investments: Evidence from natural within-family variation." *Unpublished manuscript, Brigham Young University*.

**Reynoso, Ana.** 2024. "The impact of divorce laws on the equilibrium in the marriage market." *Journal of Political Economy*, 132(12): 4155–4204.

## Acknowledgements

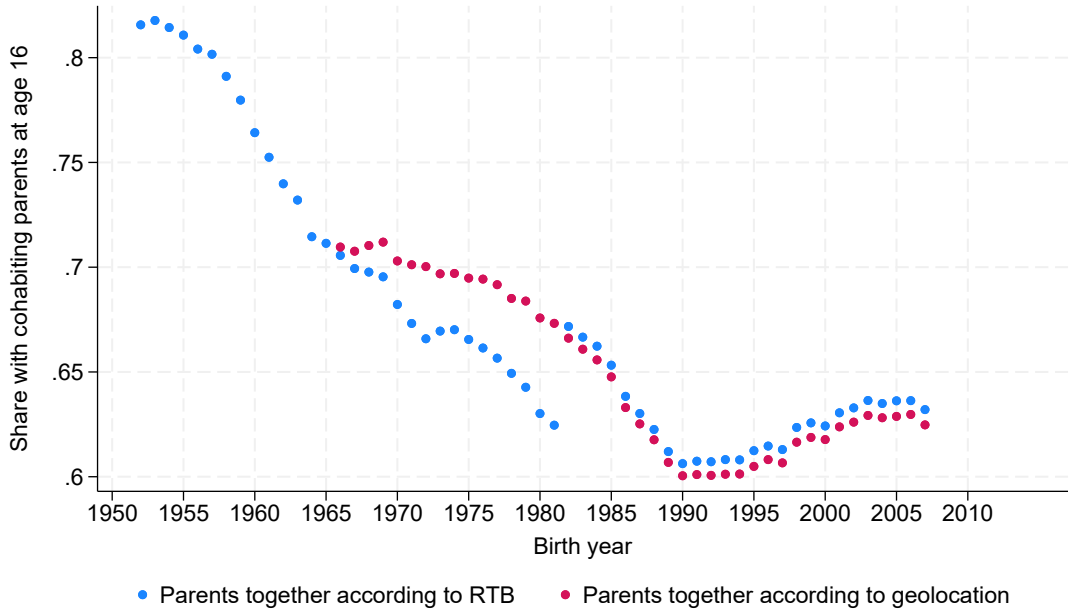
The authors have no conflicts of interest to report for this paper. Van Maarseveen gratefully acknowledges funding from the European Research Council under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. 950641).

# Appendices

## A Descriptive evidence

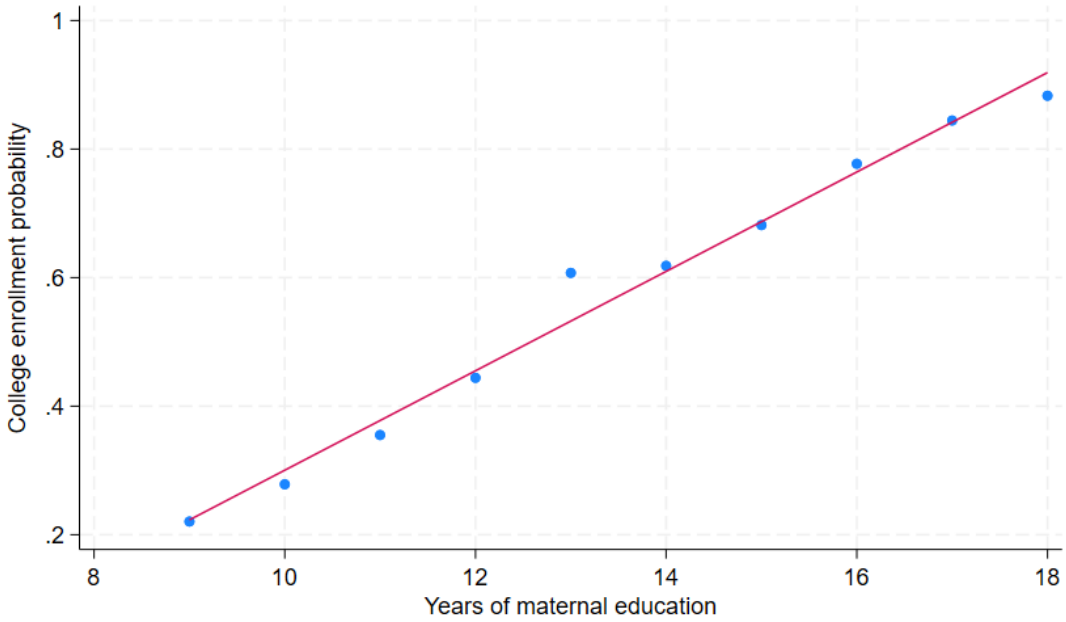
### A.1 Additional results

Figure A.1: Different sources on parental cohabitation at age 16 by birth cohort



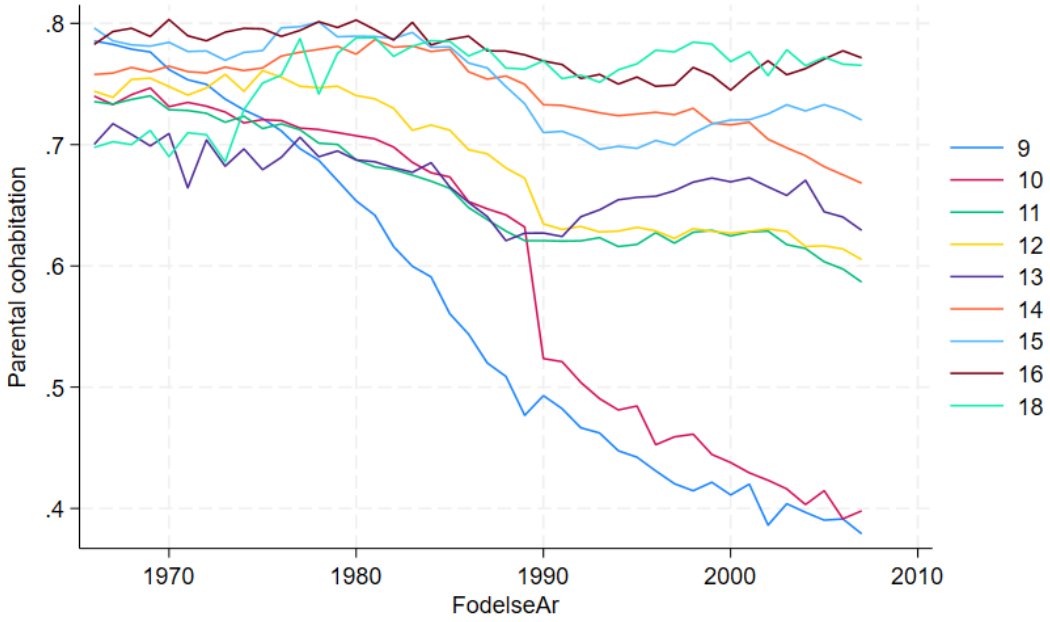

*Note*: Figure shows parental colocation probability at age 16 according to different sources. The Register of the total population begins for the cohort born in 1952 and captured legally married parents regardless of cohabitation until the cohort born in 1981. For the cohorts born in 1982-2007, the Register of the total population captures legally married parents regardless of cohabitation as well as those who cohabit. Parents are identified as cohabiting based on geographic location for the 1966-2007 cohorts if they live in the same 100x100 meter cell.

Figure A.2: College enrollment probability by maternal SES



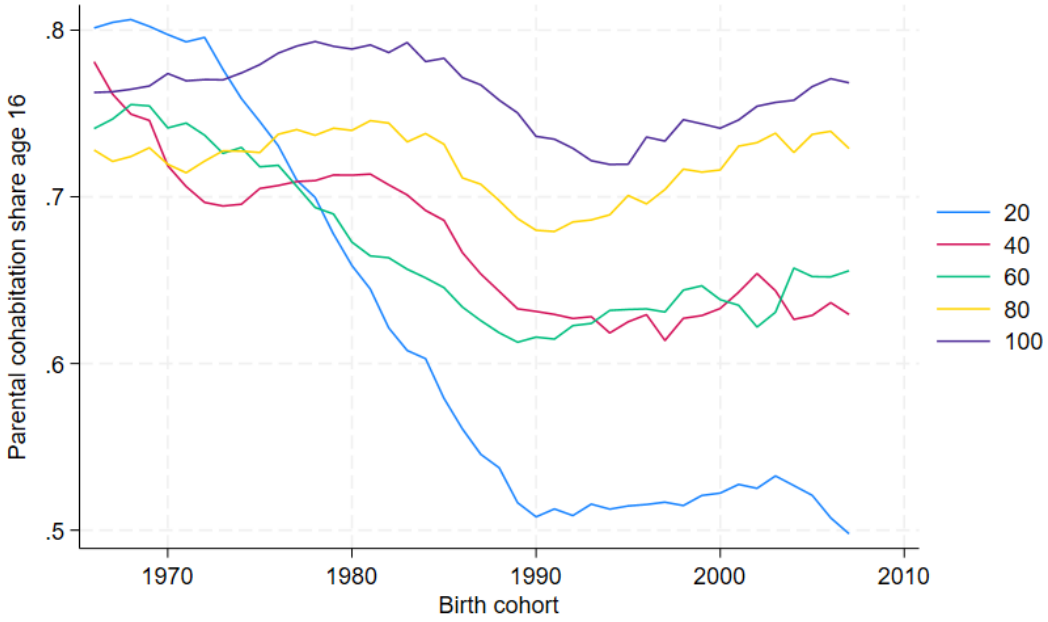

Note: Figure shows the likelihood that children register to attend a university or applied university (y-axis) by maternal years of education (x-axis). Calculations based on all children born in Sweden to Swedish-born mothers between 1995 and 2007.

Figure A.3: Parental cohabitation at age 16 by years of maternal education



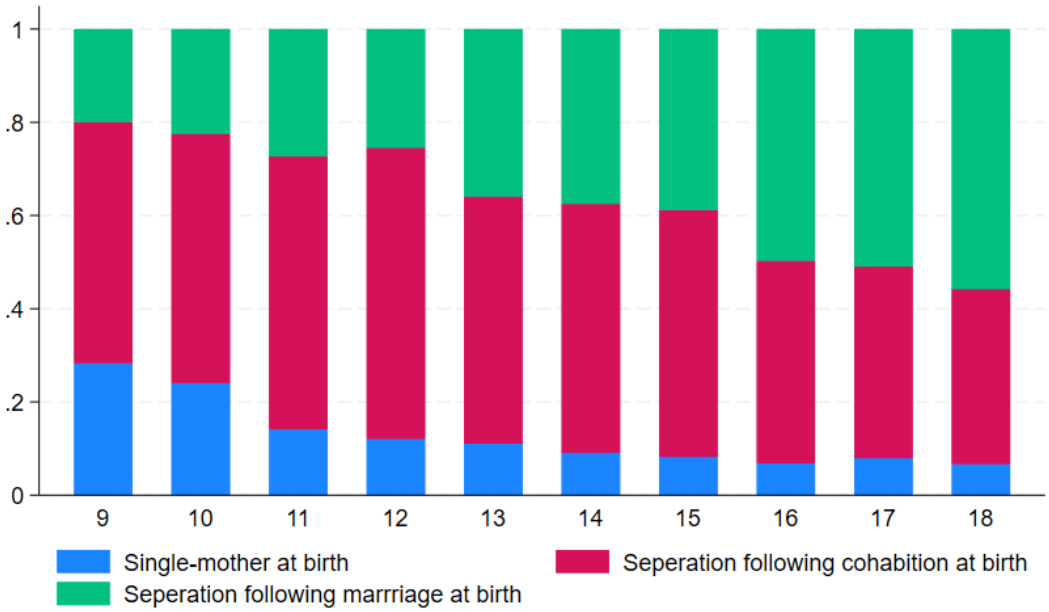

Note: Figure shows the development in parental cohabitation at age 16 by years of maternal education for the individual birth cohorts between 1966 and 2007. Note: the coding of education changes between 1999 and 2000 in the CBS LOUISE database. This due to the introduction of SunNiva2000 as the new standard, and imperfect matching between old and new educational categories. Hence, there us a large change in maternal education assignment between 1989 and 1990, since we define maternal education at 10 years after birth. Note: none of this matters for baseline estimations, since all children in the baseline sample are born in or after 1990.

Figure A.4: Development in parental cohabitation by maternal education quintile



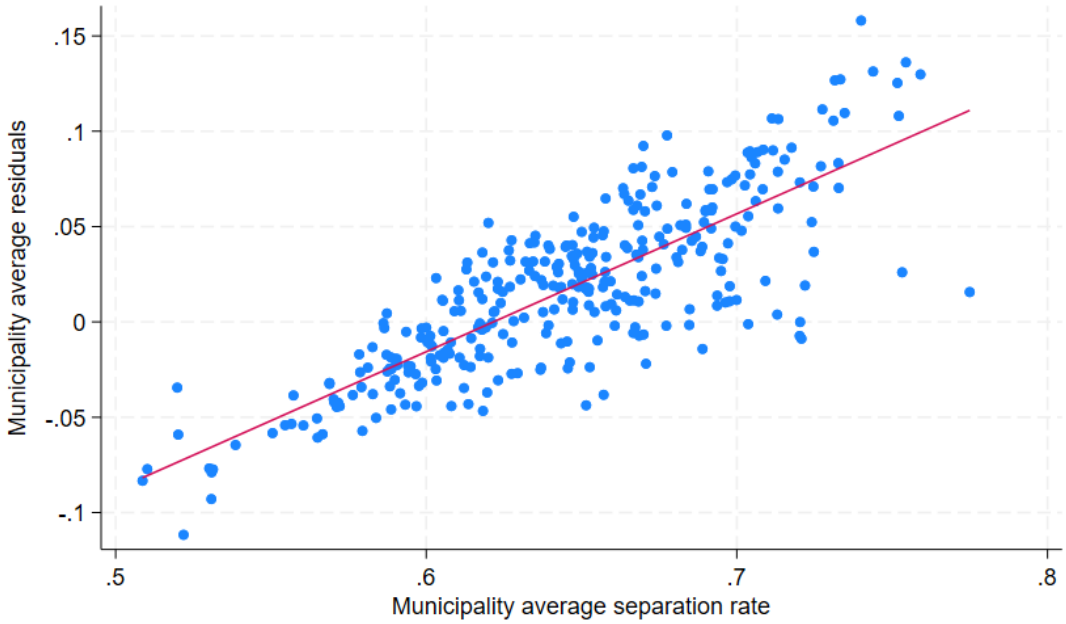

Note: Figure shows the development in parental cohabitation at age 16 by maternal education quintile. For each birth cohorts, the individuals are ordered by the 3-digit maternal education groupings (see appendix Table B.1 for more detail), and divided into 5 equal size quintiles. As such, the quintiles are based on the position of maternal education, relative to all other mothers in the same cohort. As such, each position in the graph represents 20% of children by maternal education in any given cohort.

Figure A.5: Type of separation events by maternal years of education




Note: Figure shows the type of parental separation event by years of maternal education (x-axis) for the cohorts born between 1995 - 2007. Sample restricted to children whose parents are not cohabiting at age 16.

Figure A.6: Municipality average residuals conditional on SES



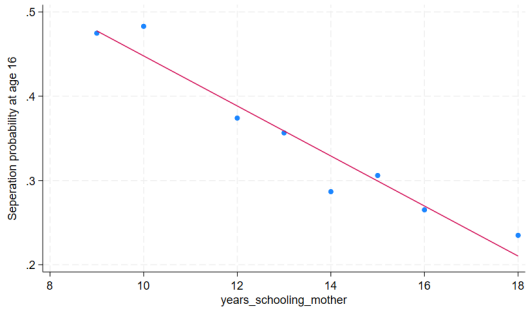

Note: Figure shows the average residuals when correcting for parental education (flexible interaction of the categories listed in Table B.1 for both parents) and age at birth on the y-axis. On the x-axis is the average separation rate by age 16 in the same municipality. The slope is 0.78, highlighting that at most 22% of the variation across space can be explained based on parental education and differences in parental age at birth.

## A.2 Inclusion of children with a migration background

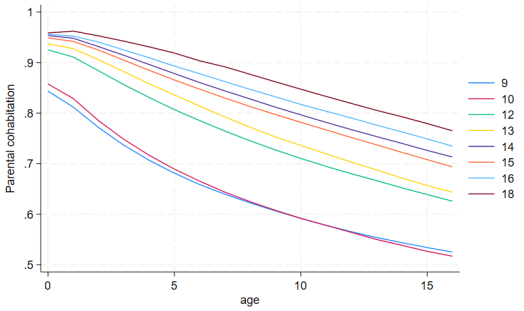
Around 16.5% of the children born between 1990 and 2000 in Sweden had a mother who was not Swedish-born. These are excluded in the baseline study, since the share of foreign-born mothers changed substantially over the study period (rising from 8% for the 1974 birth cohort to 18% in the 2001 birth cohort), and maternal education is typically less-well recorded for mothers with a migration background. Furthermore, restriction attention to the children born to Swedish-born mothers enables us to use information on the grandparents as well in the overwhelming majority of cases.

To show that the results are robust when estimated on the full population of children born in Sweden, the figures below show the key estimates when also including children with a migration background. The estimates are generally very similar to the baseline estimates, although the SES-gradient in panel (a) of figure A.7 is slightly flatter, which might be related to the higher degree of noise in educational attainment observed for foreign-born mothers.

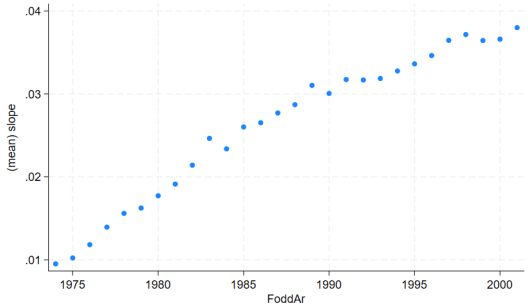
Figure A.7: SES gradient in parental separations including children with a non-Swedish born mother




(a) SES gradient in separations by age 16




(b) SES and parental colocation by age



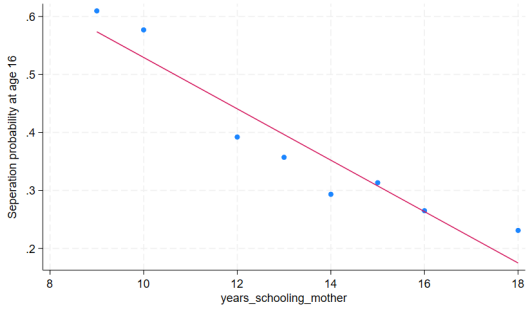

(c) Developments of SES gradient in parental separations over time

Note: Figure shows parental separation probability at age 16 by years of maternal education (left), parental co-location probability by age by years of maternal education (right), as well as development in SES-gradient over time (bottom). The figures also include children with a non-Swedish born mother.

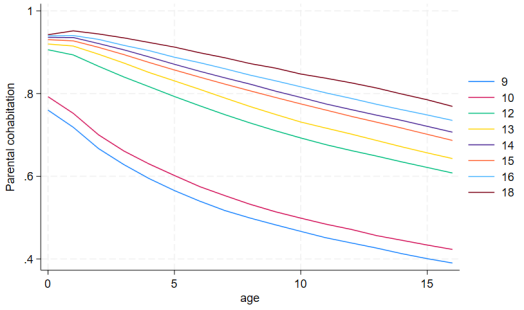
## A.3 Sample restricted to first-born children

Throughout the descriptive analyzes, we treat the child as the unit of observation. However, we classify households by SES. This may introduce some friction, since both household size and the age of siblings may certain changes over time. To show that the same results hold when providing the same weight to all households, we repeat the key descriptive analyzes based on the first-born child to each mother. This also has the benefit that if separation events are related to the age of the oldest child, we would be able to see this in different hazard rate patterns. Figure A.8 shows the key descriptive figures when restricting attention to first-born children, with very similar results in all cases.

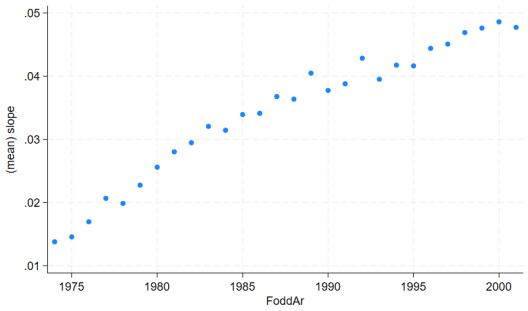
Figure A.8: SES gradient in parental separations among first-born children




(a) SES gradient in separations by age 16




(b) SES and parental colocation by age




(c) Developments of SES gradient in parental separations over time

Note: Figure shows parental separation probability at age 16 by years of maternal education (left), parental co-location probability by age by years of maternal education (right), as well as development in SES-gradient over time (bottom). All three figures are exclusively based on first-born child of the mother.

# B Educational variables

## B.1 Parental education

Information on the highest level of completed education at the individual level is available in the Swedish registers universally since 1990 via the LOUISE and LISA registers. There have been some minor changes in the allocation of individual degrees to level of completion at Statistics Sweden, in particular between 1999 and 2000. To ensure continuity across birth cohorts, we define parental education as the highest level of education observed in the registers in the most year for which information on the individual is available.

For parents that pass away prior to 1990, educational achievement is never documented in the registers. However, this forms a limited problem, given that the baseline estimates are based on children born since 1990, whose parents would thus always be captured by the relevant registers. For the very earliest cohorts the survivor bias may play some role, but this is relatively unlikely to explain any of the divergence observed since 1970, since we would observe educational attainment by the latest when the children are aged 20.

Table B.1 shows the aggregation level at which parental education is known. Within the Swedish structure, there is a clear correspondence between formal length of the education and the level of education, since vocation degrees at both the secondary and tertiary level have historically been a year shorter than academic degrees. In the reporting of results, we typically rely on the years of education. However, any regression which controls for parental education always controls for the unique combination of the educational codes for both mother and father.

Table B.1: Parental education categories

| Years | Code | Description | Share |
|---|---|---|---|
| | 100 | Elementary school education | 0.09% |
| | 106 | Elementary school education | 0.12% |
| 9 | 200 | Other primary and lower secondary education, 9 (10) years | 0.18% |
| | 204 | Lower secondary school education | 0.01% |
| | 206 | Compulsory school education, years 7–9 (10) | 6.89% |
| | 310 | Other upper secondary education, less than 2 years | 0.05% |
| | 312 | Upper secondary education, preparation studies, no diploma, less than 2 years | 0.67% |
| 10 | 313 | Upper secondary education, vocationally oriented programme, no diploma, less than 2 years | 0.63% |
| | 316 | Upper secondary education, preparation studies, less 2 years | 0.00% |
| | 317 | Upper secondary education, vocationally oriented, less than 2 years | 2.76% |
| | 320 | Other upper secondary education, 2 years | 0.41% |
| | 322 | Upper secondary education, preparation studies, no diploma, 2 years | 0.60% |
| 11 | 323 | Upper secondary education, vocationally oriented, no diploma, 2 years | 1.01% |
| | 326 | Upper secondary education, preparation studies, 2 years | 3.55% |
| | 327 | Upper secondary education, vocationally oriented, 2 years | 16.66% |
| | 330 | Other upper secondary education, 3 years | 0.25% |
| | 332 | Upper secondary education, preparation studies, no diploma, 3 years | 1.22% |
| 12 | 333 | Upper secondary education, vocationally oriented, no diploma, 3 years | 5.08% |
| | 336 | Upper secondary education, preparation studies, 3 years | 9.04% |
| | 337 | Upper secondary education, vocationally oriented, 3 years | 6.93% |
| | 410 | Other post-secondary education, less than 2 years | 0.01% |
| | 412 | University/college programme, minimum 30 higher education credits, no degree, less than 2 years | 3.66% |
| 13 | 413 | Upper secondary supplementary programme, less than 2 years | 1.10% |
| | 415 | Post-secondary education, not university/college, vocationally oriented, less than 2 years | 0.73% |
| | 417 | University/college programme, vocationally oriented, less than 2 years | 0.25% |
| | 520 | Other post-secondary education, 2 years | 0.07% |
| | 522 | University/college programme, 120 higher education credits, no degree, 2 years | 1.50% |
| 14 | 525 | Post-secondary education, not university/college, vocationally oriented, 2 years | 1.95% |
| | 526 | University/college programme, general, 2 years | 0.43% |
| | 527 | University/college programme, vocationally oriented, 2 years | 7.17% |
| | 530 | Other post-secondary education, 3 years | 0.01% |
| | 532 | University/college programme, minimum 180 higher education credits, no degree, 3 years | 2.28% |
| 15 | 535 | Post-secondary education, not university/college, professionally oriented, 3 years | 0.07% |
| | 536 | University/college programme, general, 3 years | 4.40% |
| | 537 | University/college programme, professionally oriented, 3 years | 10.16% |
| | 540 | Other post-secondary education, 4 years | 0.00% |
| 16 | 545 | Post-secondary education, not university/college, professionally oriented, 4 years | 0.04% |
| | 546 | University/college programme, general, 4 years | 1.93% |
| | 547 | University/college programme, professionally oriented, 4 years | 5.78% |
| | 550 | Other post-secondary education, 5 years or longer | 0.00% |
| 17 | 555 | Post-secondary education, not university/college, professionally oriented, 5 years | 0.02% |
| | 556 | University/college programme, general, 5 years or longer | 0.02% |
| | 557 | University/college programme, professionally oriented, 5 years or longer | 1.30% |
| | 600 | Other postgraduate education | 0.00% |
| 18 | 620 | Licentiate programme | 0.13% |
| | 640 | Doctoral programme | 0.74% |
| - | 999 | Missing | 0.05% |
| | Total | | 100.00% |

Note: Figure shows the links between years of education and underlying codes. Shares based on distribution of maternal education for children born in Sweden between 1990 - 2007.

## B.2 Educational outcomes of children

The educational outcomes consist of three key measures: the scores on the national standardized tests, the primary school GPA, and enrollment and/or graduation

from the academic track. We will discuss each in some detail, particularly the standardized tests.

### B.2.1 National standardized tests

Sweden has historically relied on national standardized tests to provide insights into the educational development of children. Since 1998, Sweden has conducted a national test in grade 9, which is typically taken at the end of compulsory education in the year in which children turn 16. The grade 9 standardized tests results are available as part of the Statistics Sweden (SCB) databases since 2003, thus corresponding to the 1987 birth cohort. However, this provided only limited insights into the human capital development of children prior to age 15 for the purposes of public policy development, monitoring, and interventions. As such, an additional national test was introduced in grade 3 in 2010 (available for the birth cohorts starting in 2000), and a further national test was introduced in grade 6 in 2012 (available for the birth cohorts starting in 1999). Testing was suspended during COVID, which means that national tests are not available for 2020 and 2021.

Each national tests covers students their ability in Swedish and English. The grade 6 and 9 standardized tests in addition tests for the English ability. There are certain additional subjects that rotate or that have been available for only certain years. We focus on the scores in the three core subjects of English, Mathematics, and Swedish, to enhance availability over time and between different cohorts.

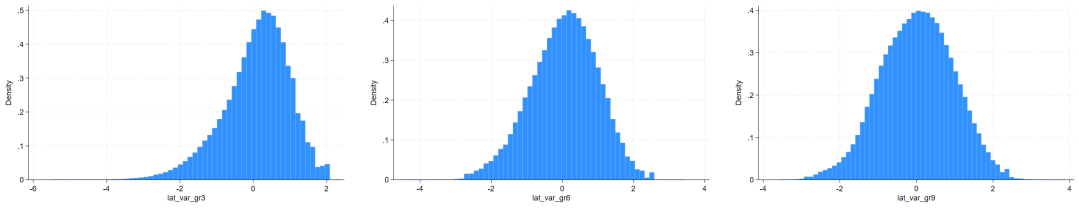
It is difficult to directly transform the scores into a single aggregate score, since the individual subjects typically have scores across a variety of sub-domains or sub-test parts. These scores can be binary (pass/fail), letter grade (A-F), or a numerical grade (number of correct questions). A combination of all three can often be found as part of the assessments of a single subject in a single year (say Swedish in 2013). As such, this creates a complication, since it's difficult to transform the letter grades and pass/fails into a numerical grade, given that each measure has a clear ordinal meaning, but not a clear cardinal interpretation.

To obtain a comparable ability measure that uses as much information as possi-

ble, we estimate a latent factor model for each test-year. We combine the grades in Swedish, Mathematics, and English into a single ability measure to maximize precision. This also mimics the way in which the primary school GPA is constructed, which is the average grade across 16 subjects, which forms the key input into secondary school decisions. Given the ordinal nature of the variables, we estimate an item response theory model at the individual level. These have been widely used in economics in recent years in other publications on early childhood measures with similar structured data, typically from surveys (see for instance Andrew et al. (2024); Bau (2022); Carneiro et al. (2024); Kinsler and Pavan (2021)). However, the results are similar when using an average of the numerical scores in mathematics, Swedish, and English and the correlations are fairly high at 0.90 - 0.93.

The inferred latent ability measures follow quite well a standard distribution, as shown in Figure B.9. The benefit of the rich data and variety of tests is that top-coding is not much of an issue, as there is variation even in the top of the ability distribution in the number of correct items. Furthermore, the ability tests are highly predictive, as shown in Figure B.10 for the grade 3 ability. Children at the bottom of the ability distribution have a probability of less than 20% to attend the academic track, vs over 90% for children at the top of the grade 3 ability distribution. These relationships become substantially stronger over time as seen for instance based on the relationship between the grade 6 national test and later educational outcomes, as shown in Figure B.11.

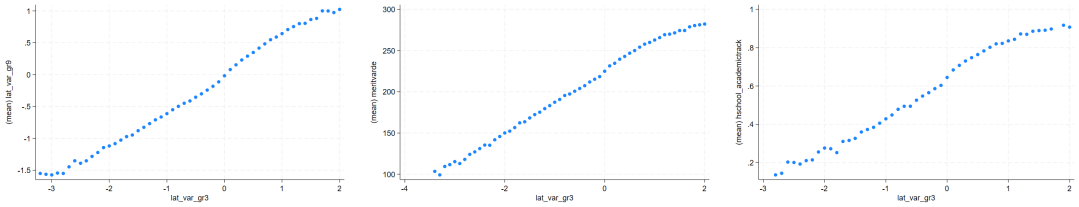
Figure B.9: Distribution of latent ability measures based on test scores




(a) Grade 3 (b) Grade 6 (c) Grade 9

Note: Distribution of latent ability based on grade 3, 6, and 9 test scores for all cohorts with available test data.

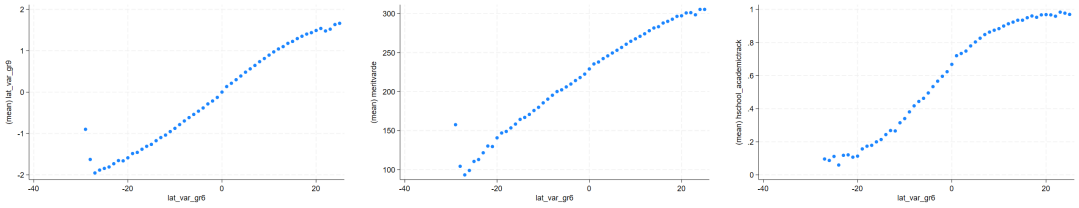
Figure B.10: Predictiveness of latent ability measures - grade 3



(a) Grade 9 test score (b) Primary school GPA (c) Academic HS Track

Note: Figure shows the relationship between grade 3 latent ability and 9th grade latent ability, primary school GPA, and academic track completion in high school.

Figure B.11: Predictiveness of latent ability measures - grade 6



(a) Grade 9 test score (b) Primary school GPA (c) Academic HS Track

Note: Figure shows the relationship between grade 6 latent ability and 9th grade latent ability, primary school GPA, and academic track completion in high school.

### B.2.2 Primary school GPA

The primary school GPA is determined at the end of primary school. The grade is determined as a weighted average of the grades on the underlying 16 or 17 subjects in which students graduated.[8] The primary school grades are assigned by the teachers at the end of compulsory education. The primary school GPA forms the key determinant in the choices for secondary schools, since the municipality assign students to their preferred secondary school tracks in order of their primary school GPA. As such, the primary school GPA is the key determinant of secondary school educational opportunities.

The primary school GPA roughly follows a normal distribution. To improve interpretation, we standardize the primary school GPA within each graduation cohort

[8]The underlying subjects are fixed and common for all students, and include Swedish, English, Mathematics, Biology, Physics, Chemistry, History, Geography, Religion, Civics, PE, Crafts, Music, Visual arts, Home and Consumer Studies, Technology, and a second foreign language.

to have mean zero and SD 1, which improves the comparison with the other outcomes.

### B.2.3 Academic track enrollment/graduation

Records on the academic track graduation are available at the student-level since 1973, although the full set of information is available since 1985. Throughout the paper, we focus on whether students graduate from an academic high school track since this is best documented in the data throughout the entire period. In recent years, around 60% of students graduates from the academic track in high school. The grades in the academic track are the key admission criteria into tertiary education, and $2/3^{\text{rd}}$ of spots in each study program is assigned based on high school GPA. The other $1/3^{\text{rd}}$ or spots are allocated on the basis of the score on the Swedish Scholastic Aptitude Test (SSAT), the so-called Högskoleprovet, which is comparable to the SATs in the US context. As such, gaining access to a high-quality academic track is a key input into future academic outcomes, although the second chance does always exists through participation in the SSAT.

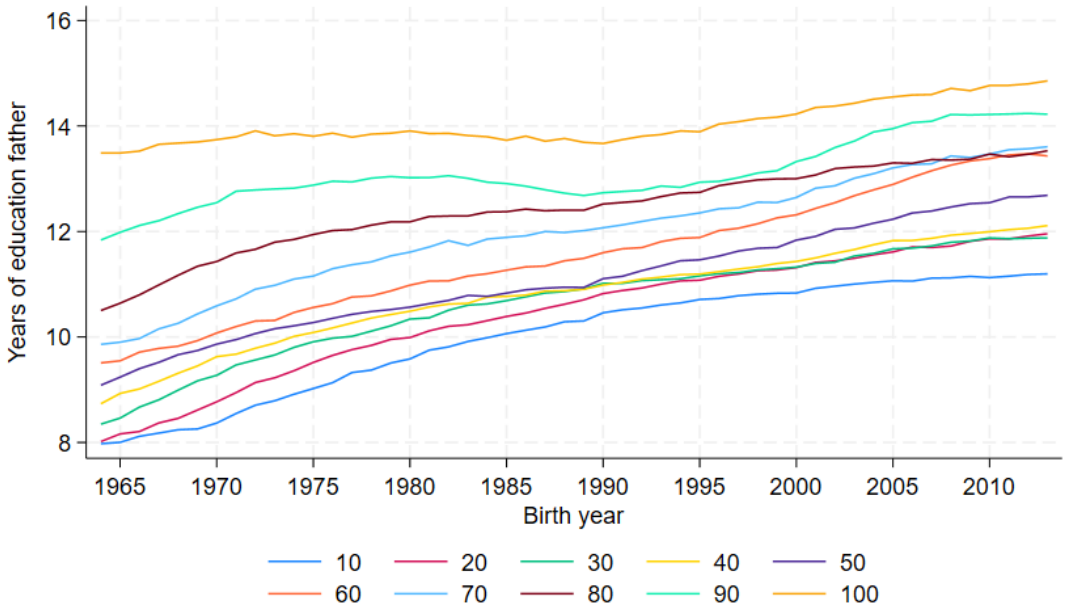
# C Potential mechanisms

## C.1 Assortative matching

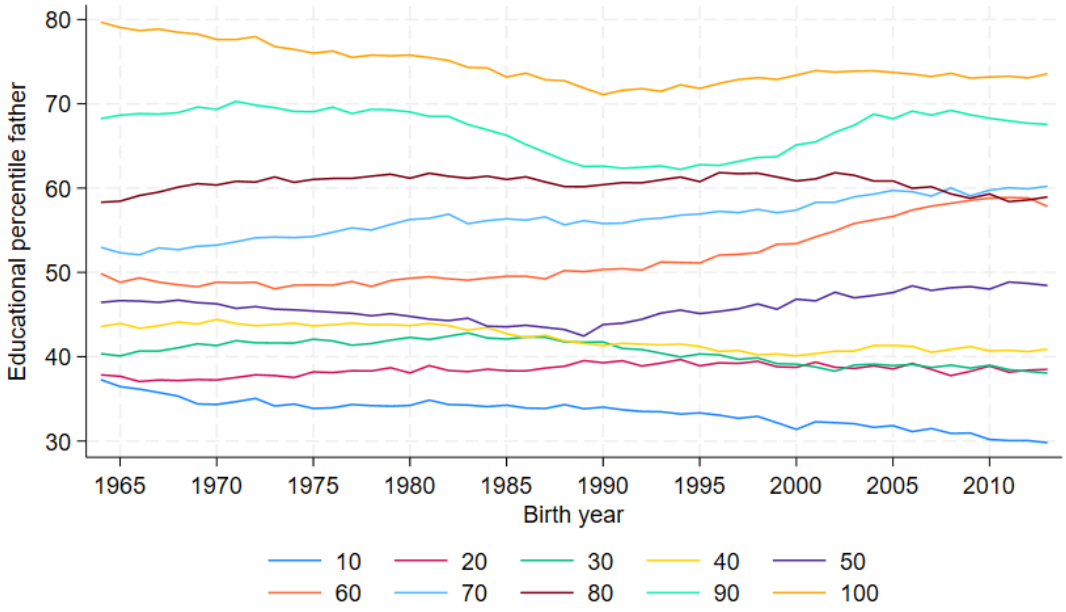
One possibility is that the decline in relationship stability among children born to low-educated mothers is due to changes in partner selection. If females married to low-SES partners are generally more likely to separate, and assortative matching substantially increased, then this may reasonably drive the increase in separation rates. To investigate this, panel (a) in Figure C.12 below shows how the educational position of fathers (as percentile in the educational distribution) has changed between 1964 and 2013. While there has been some increase in assortative matching for females in the bottom 10% of the educational distribution, the patterns are generally quite stable. This is remarkably, since we observed a decline in family stability for females in the bottom 60% of the educational distribution (see appendix Figure A.4). As such, the changes are not quite consistent with a change in assortative matching defined by educational percentiles.

The results are even more striking when studying the changes in assortative matching in levels, as highlighted in panel (b) of Figure C.12. While the average years of education of the father has barely changed over a 50 year period for females in the top-decline of the educational distribution, mothers in the bottom educational decile now have partners that have received over 3 years of additional education compared to the 1960's. As such, the inequality in partners between females of different educational levels has actually decreased when education is measured in years of education, and remained relatively stable in terms of the educational position of the fathers. As such, there is little evidence of a substantial increase in assortative matching over the period under study.

Figure C.12: Changes in assortative matching by maternal education




(a) Father's years of education




(b) Father's educational percentile

Note: Figure shows the change in assortative matching based on the father's education by the maternal education decile between 1964 - 2013. Panel (a) shows the father's percentile position within all fathers that have children in the same year. Panel (b) shows years of education of the father.

# D Effects of parental separations

## D.1 Birth spacing in HH FE models

One complication factor in the estimation of the household-fixed effect models is controlling for the birth-spacing of children. Given that within a household the difference in age at birth is fully co-linear with the difference in age at time of separation, it's impossible to estimate these separately. To see this, imagine that we estimate the simple estimating model below on households with two children, where we include a continuous birth year to account for educational trends over time:

$$\text{Educational outcome}_{ih} = \alpha_h + \beta_1 \text{Age at separation}_i + \beta_2 \text{Birth year}_i + \Gamma X_i + \epsilon_i \tag{2}$$

Given the inclusion of the household fixed effect, the model is equivalent to taking first differences. As such, the estimating equation based on children $i$ and $j$ in household h is equal to:

$$\Delta\text{Educational outcome}_{ij} = \beta_1 \Delta\text{Age at separation}_{ij} + \beta_2 \Delta\text{Birth year}_{ij} + \Gamma \Delta X_{ij} + \Delta\epsilon_{ij} \tag{3}$$

Since the time of the separation event is identical for both siblings, the difference in age at separation is equal to the birth spacing (i.e., $\Delta\text{Age at separation}_{ij} = -\Delta\text{Birth year}_{ij}$) and their relative effects cannot be independently estimated.

In our context, controlling for birth years as such should be relatively unimportant, since the educational attainment variables are normalized within each birth cohort (see appendix B.2 for more detail). If the difference in birth year between siblings (i.e., the birth-spacing) is also unimportant for educational attainment, then the fact that we cannot independently estimate the effects of birth spacing from the effect of difference in age at time of separation would be comparatively unimportant.

However, this is far from the case in practice. To estimate the effect of birth-spacing and birth-order, we focus on households with two children for simplicity. Table D.2 below shows the effect of birth-order and birth-spacing on the scores on

the national grade 3, 6, and 9 ability tests.

As can be seen from Table D.2, first-born children generally perform better compared to second-born children. However, the difference depends on the birth-spacing between children. The larger the difference in birth-spacing, the smaller the difference in ability between children on the various national tests. Generally, a increase in birth-spacing by 1 year decreases the ability of the first-born child relative to the second-born child by 0.015 - 0.02 SD.

There are many factors that may be at play here, including competition between children for limited parental resources and crowding-out. Alternatively, the timing of birth-spacing is endogenous to the ability of the first child.[9] Previous research already suggested that birth spacing may not be ortogonal to the realized ability of children, as highlighted by Price and Kalil (2010) and Buckles and Munnich (2012).[10] However, this creates a difficulty, since the difference in age at separation between siblings is fully co-linear with birth spacing between siblings, as highlighted above. As such, the effects of differences in the age at separation would also capture the difference in ability of siblings related to birth spacing.

One solution is to include the households whose families do not separate to correct for the effects of birth-spacing. The key assumption is that the relationship between birth-spacing and potential ability of siblings is similar between households who do, and do not, separate during childhood. Table D.3 below shows the relationship between birth-spacing and relative ability of siblings both within households do and do not separate during childhood. The effect of birth spacing on the relative ability is similar across both samples, highlighting that using non-separating parents may be good way to control for the effects of birth-spacing. In this case, we can correctly identify the effects of experiencing separation events at various ages from the comparison of siblings in separated household conditional on the effects of birth spacing. This is the approach we follow in the main estimations

[9] One can for instance imagine that complications during pregnancy, birth, or early-childhood affects the subsequent birth-spacing decisions by parents.

[10] Interestingly, we find that an increase is birth spacing by a year decreases the difference in ability between the oldest and youngest child by 0.015 - 0.02 SD, whereas Buckles and Munnich (2012) find that it increases the difference by 0.017 SD. The literature on the effects of birth-spacing is fairly small, meaning that few estimates are generally available.

Table D.2: Effects of parental separation on educational outcomes

| | Grade 3 test | Grade 6 test | Grade 9 test |
|---|---|---|---|
| | (1) | (2) | (3) |
| First Born | 0.113*** | 0.100*** | 0.186*** |
| | (0.00974) | (0.0117) | (0.00810) |
| Birth spacing | -0.0156*** | -0.0198*** | -0.0212*** |
| | (0.00295) | (0.00358) | (0.00233) |
| $N$ | 246510 | 160132 | 279426 |
| $R^2$ | 0.664 | 0.720 | 0.743 |

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Estimate based on all households where two children have available information on the national tests. First born refers to the older of the two siblings. Birth spacing is defined as the difference in birth years. Sample restricted to children where both the father and mother had no children from a previous relationship.

in order to correct for the effects of birth-spacing.

Table D.3: Effects of parental separation on educational outcomes

| | Grade 3 test | | Grade 6 test | | Grade 9 test | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| First Born | 0.113*** | 0.125*** | 0.100*** | 0.0887*** | 0.186*** | 0.193*** |
| | (0.00974) | (0.0147) | (0.0117) | (0.0169) | (0.00810) | (0.0104) |
| Birth spacing | -0.0156*** | -0.0188*** | -0.0198*** | -0.0161** | -0.0212*** | -0.0241*** |
| | (0.00295) | (0.00464) | (0.00358) | (0.00544) | (0.00233) | (0.00318) |
| $N$ | 246510 | 109550 | 160132 | 75510 | 279426 | 180212 |
| $R^2$ | 0.664 | 0.665 | 0.720 | 0.723 | 0.743 | 0.741 |
| Parents separated in childhood? | No | Yes | No | Yes | No | Yes |

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Estimate based on all households where two children have available information on the national test scores. First born refers to the older of the two siblings. Birth spacing is defined as the difference in birth years on the youngest child, relative to the oldest. In both cases, sample restricted to children where both the father and mother had no children from previous relationships.

## D.2 Parameterized effects on educational outcomes

Estimating equation 1 provides estimates for each single age at separation. While the data generally allows for the estimation of the individual effects, these are less suitable for comparison between groups and robustness checks. To simplify comparisons between different models, we put some structure on individual $b_m$ estimates. Specifically, we group the $\beta_m$ estimates in age at separation bands as shown in equation 4. The omitted group remains children whose families separated when they were between ages 18 - 20. As before, a household fixed effect ($\alpha_h$) is included so estimations are based on comparisons within households.

$$\text{Educational outcome}_{ihc} = \alpha_h + \beta_{1-4} I(m_i \leq 4) + \beta_{5-8} I(5 \leq m_i \leq 8) + \beta_{9-12} I(9 \leq m_i \leq 12) + \beta_{13-15} I(13 \leq m_i \leq 15) + \beta_{16-17} I(16 \leq m_i \leq 17) + \gamma_c + \Gamma X_i + \epsilon_i \tag{4}$$

In this case, the $\beta_{1-4}$ estimates capture the effect of separation during the first four years of life, $\beta_{5-8}$ the effect of separation during the years before the grade 3 national test, $\beta_{9-12}$ the effect of separations during the years before the grade 6 national test, $\beta_{13-15}$ the effect of separation during the years just before finishing compulsory education, and $\beta_{16-17}$ the effects of separation in the early high-school years. The omitted category are ages 18 - 20, when children are typically finishing high-school and are deciding whether to enroll in tertiary education.

To illustrate the effects and provide a baseline for comparison, Table D.4 shows the $\beta$ estimates obtained from estimation equation 4 on the baseline sample. The sample is the same as that used to estimate the individual $\beta_m$ estimates in Figure 12. The results in Table D.4 show similar effects to those derived from the main Figure. For the three national tests, parental separation events have the largest negative effects when they occur in the years prior to the measurement of the outcome. As such, it appears that the disruption effects dominate and potential length-of-exposure to separation effects.

Interestingly, we again see that for the primary school GPA, the key determinant is whether the separation event occurs at any time prior to finishing compulsory

Table D.4: Effects of parental separation on educational outcomes

| | (1) Grade 3 test | (2) Grade 6 test | (3) Grade 9 test | (4) Grade 9 GPA | (5) HS Track |
|---|---|---|---|---|---|
| Separated age 1 - 4 | -0.0343 | -0.0175 | 0.0117 | -0.120*** | -0.0215* |
| | (0.0207) | (0.0249) | (0.0150) | (0.0144) | (0.00959) |
| Separated age 5 - 8 | -0.0506** | -0.0369 | -0.0174 | -0.106*** | -0.0229** |
| | (0.0183) | (0.0218) | (0.0130) | (0.0125) | (0.00827) |
| Separated age 9 - 12 | -0.0417* | -0.0607** | -0.0254* | -0.117*** | -0.0213** |
| | (0.0162) | (0.0189) | (0.0112) | (0.0107) | (0.00706) |
| Separated age 13 - 15 | -0.0254 | -0.0263 | -0.0465*** | -0.130*** | -0.0295*** |
| | (0.0145) | (0.0162) | (0.00963) | (0.00916) | (0.00598) |
| Separated age 16 - 17 | -0.00257 | -0.0166 | -0.00688 | -0.0397*** | -0.00679 |
| | (0.0139) | (0.0148) | (0.00902) | (0.00835) | (0.00555) |
| $N$ | 477401 | 297076 | 595563 | 615336 | 542740 |
| $R^2$ | 0.647 | 0.716 | 0.726 | 0.738 | 0.666 |

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Estimated effects obtained by estimation equation 4 on national standardized test scores, primary school GPA, and high school academic track completion. The baseline group is those whose parents separate when the child is aged 18 - 20. Household-fixed effects are included, so estimates are obtained from sibling comparisons. Sample restricted to first children born to both parents. Standard errors are cluster at the household level.

education. Separating between ages 1 - 4, 5 - 8, 9 - 12, and 13 - 15 all result in roughly a similar decline in GPA by 0.11 - 0.13 SD compared to a parental separation event at ages 18 - 20. The same pattern holds for the choice of high school track (academic or vocational). As such, while the disruption effects dominate for the national test scores, it's the fact whether parents are separated or not by the end of primary school that drives the effect on primary school GPA and subsequent academic track completion.

## D.3 Additional robustness checks and heterogeneity

Table D.5: Effects of parental separation on educational outcomes - maternal education more than high school

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| | Grade 3 test | Grade 6 test | Grade 9 test | Grade 9 GPA | HS Track |
| Separated age 1 - 4 | -0.0355 | -0.0217 | 0.0179 | -0.0850*** | -0.0291* |
| | (0.0272) | (0.0337) | (0.0235) | (0.0218) | (0.0139) |
| Separated age 5 - 8 | -0.0602* | -0.0331 | -0.0183 | -0.0773*** | -0.0284* |
| | (0.0240) | (0.0291) | (0.0199) | (0.0185) | (0.0116) |
| Separated age 9 - 12 | -0.0588** | -0.0525* | -0.0333* | -0.0991*** | -0.0312** |
| | (0.0212) | (0.0251) | (0.0169) | (0.0157) | (0.00984) |
| Separated age 13 - 15 | -0.0457* | -0.0228 | -0.0408** | -0.104*** | -0.0329*** |
| | (0.0188) | (0.0216) | (0.0144) | (0.0133) | (0.00826) |
| Separated age 16 - 17 | -0.0194 | -0.00737 | -0.00352 | -0.0302* | -0.0146 |
| | (0.0178) | (0.0197) | (0.0133) | (0.0120) | (0.00761) |
| $N$ | 290667 | 175516 | 266292 | 273340 | 249533 |
| $R^2$ | 0.622 | 0.688 | 0.697 | 0.709 | 0.624 |

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

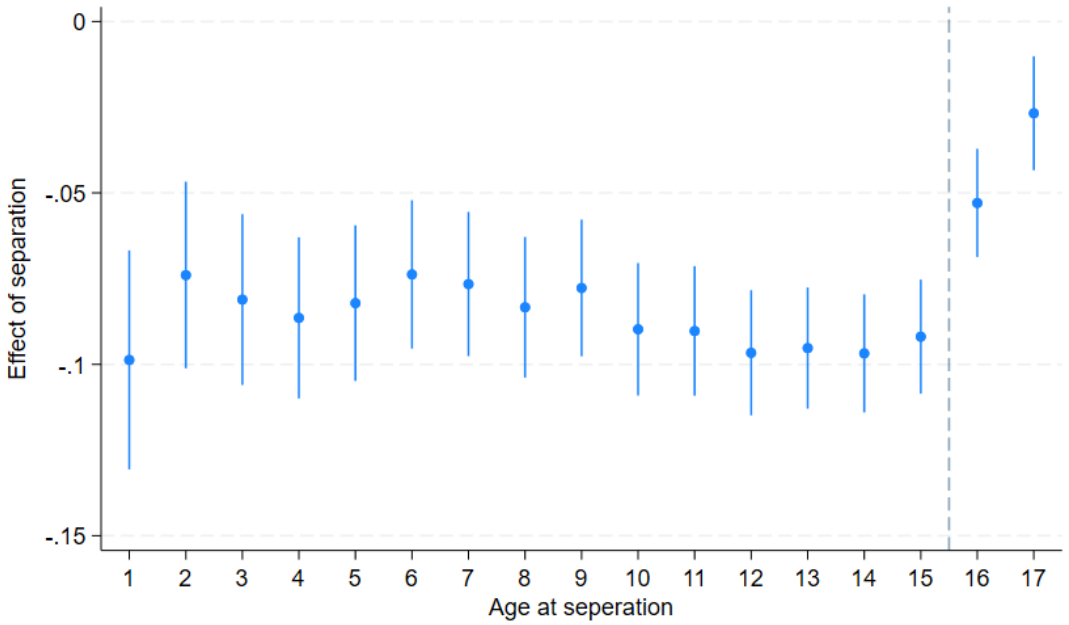
Table D.6: Effects of parental separation on educational outcomes - maternal education high school or less

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| | Grade 3 test | Grade 6 test | Grade 9 test | Grade 9 GPA | HS Track |
| Separated age 1 - 4 | -0.0220 | -0.0344 | -0.00778 | -0.145*** | -0.0151 |
| | (0.0332) | (0.0388) | (0.0206) | (0.0200) | (0.0139) |
| Separated age 5 - 8 | -0.0332 | -0.0558 | -0.0271 | -0.128*** | -0.0191 |
| | (0.0295) | (0.0343) | (0.0180) | (0.0176) | (0.0122) |
| Separated age 9 - 12 | -0.0197 | -0.0826** | -0.0226 | -0.134*** | -0.0161 |
| | (0.0261) | (0.0299) | (0.0156) | (0.0152) | (0.0104) |
| Separated age 13 - 15 | 0.000676 | -0.0428 | -0.0467*** | -0.147*** | -0.0270** |
| | (0.0232) | (0.0255) | (0.0135) | (0.0131) | (0.00892) |
| Separated age 16 - 17 | 0.0193 | -0.0316 | -0.00800 | -0.0455*** | -0.00199 |
| | (0.0225) | (0.0231) | (0.0127) | (0.0119) | (0.00825) |
| $N$ | 177851 | 115804 | 312444 | 324592 | 277440 |
| $R^2$ | 0.633 | 0.694 | 0.696 | 0.718 | 0.644 |

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

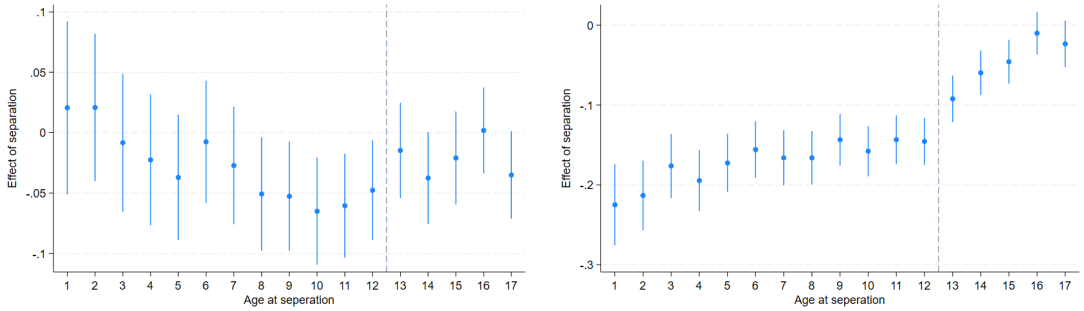
Figure D.13: Age at separation and grade 9 GPA based on English, Mathematics, and Swedish




Note: Estimated effects of age at parental separation on primary school GPA based on the same three subjects as are part of the grade 9 national tests (English, mathematics, and Swedish). The baseline group is those whose parents separate when the child is aged 18 - 20, the estimates show the effects of experiencing a separation event at age $m$ instead of age 18 - 20. Household-fixed effects are included, so estimates are obtained from sibling comparisons. The dotted lines show the age at which outcomes are measured. Standard errors are cluster at the household level.

## D.4 Extension: Grade 6 outcomes

Figure D.14: Grade 6 outcomes and age at separation (in SD)




(a) Grade 6 national test score (b) Primary school grade 6 GPA

Note: Estimated effects of age at parental separation on grade 6 standardized test score and grade 6 GPA (all normalised within birth cohorts). The omitted group is those whose parents separate when the child is aged 18 - 20, the estimates show the effects of experiencing a separation event at age $m$ instead of age 18 - 20. Household-fixed effects are included, so estimates are obtained from sibling comparisons. Sample restricted to children born to parents who did not have children with anyone else previously. Standard errors are cluster at the household level.

Table D.7: Cross-sectional difference in grade 6 outcomes (in SD)

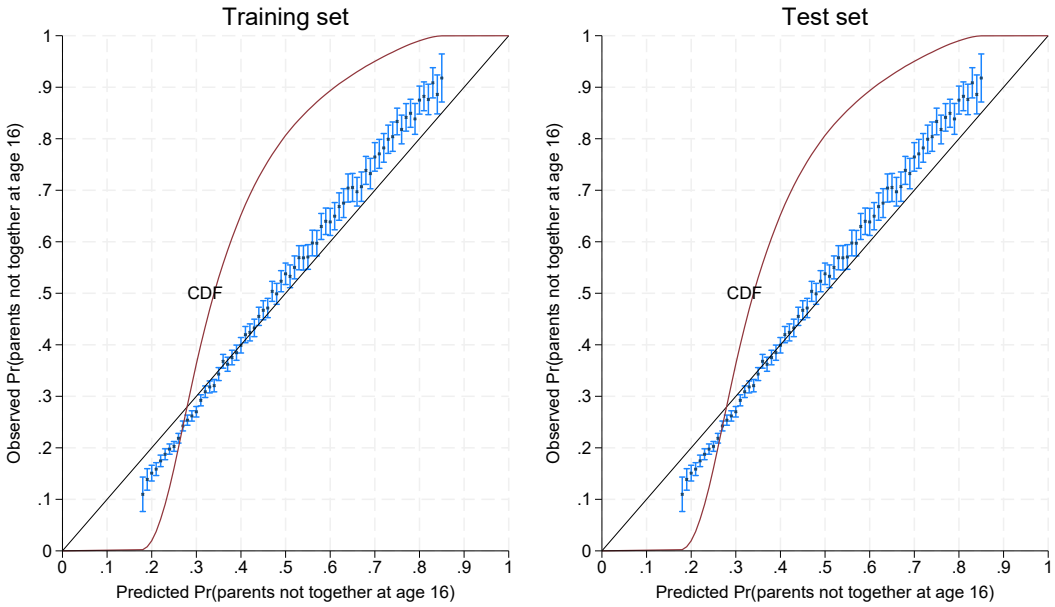
| | Grade 6 test score | | Grade 6 GPA | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Parents cohabit in year prior to test? | 0.208*** | 0.0379*** | 0.312*** | 0.156*** |
| | (0.0021) | (0.0020) | (0.00248) | (0.00235) |
| N | 552,452 | 552,452 | 549,827 | 549,827 |
| $R^2$ | 0.01 | 0.17 | 0.03 | 0.19 |
| SES controls | No | Yes | No | Yes |

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Estimates for all cohorts for which the relevant national tests scores are available. SES controls include a flexible interaction of the parental education according to the categories listed in appendix Table B.1 and controls for parental age at birth.

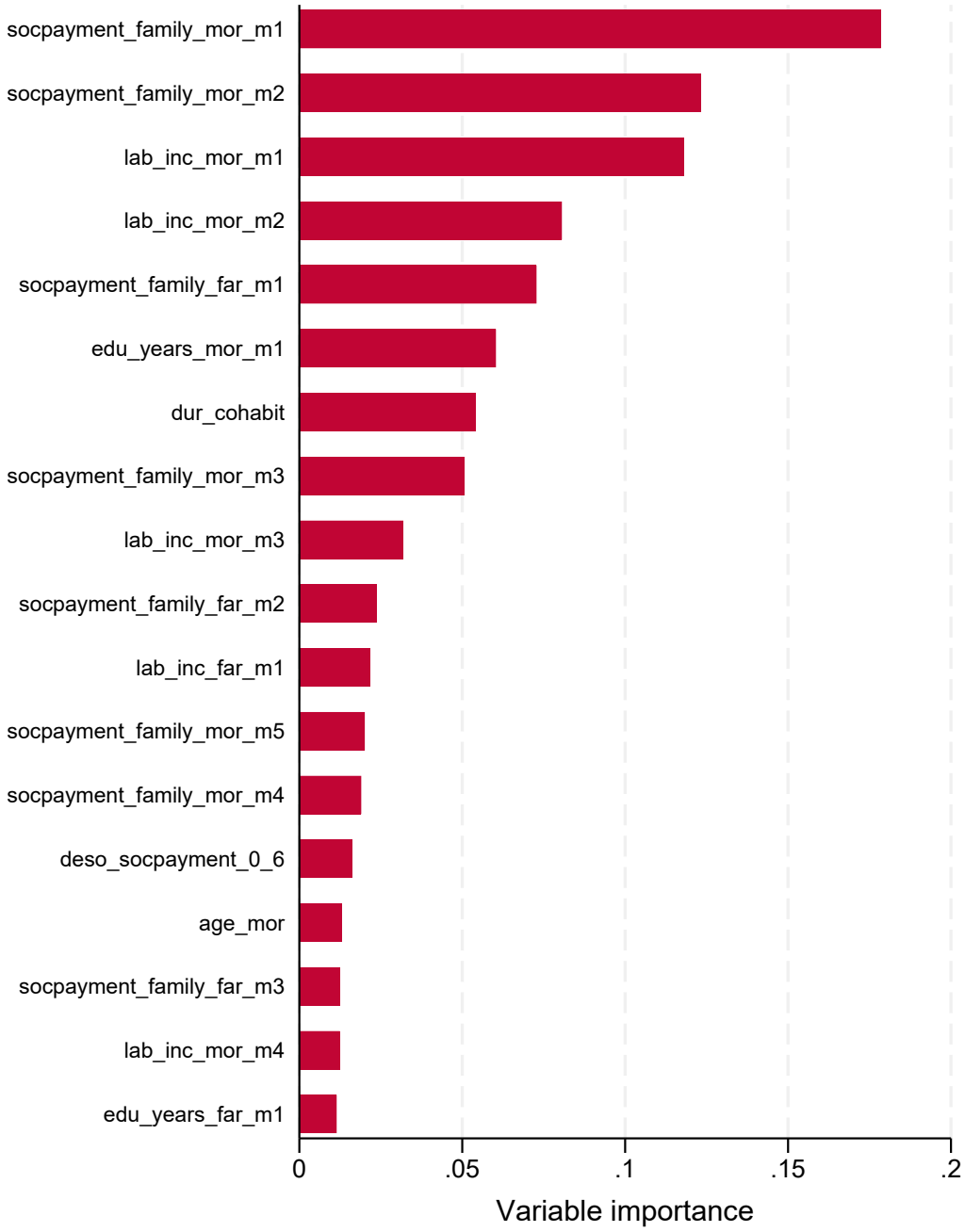
# E Generalized random forest predictions

Figure E.15: Calibration of regression forest for predicting separations




Note: Observed Pr(separation) in each percentage point of predicted Pr(separation) in training set and held-out test set. Only cells with at least 100 matches shown. CDF shows share of matches whose breakup probability is lower than a given value. 95 percent confidence intervals.

Figure E.16: Simple variable importance measure from the regression forest




Note: Bars represent share of times the regression forest has split on a particular variable in its trees at depth below 4. Splits at lower depth level given twice the weight of splits at the next depth level.