



# Correction of Bias in the Index of Family Belonging & Rejection

Henry Potrykus<sup>1</sup>

February 12, 2014

## Introduction & Statement of Work

The Index of Family Belonging and Rejection (hereafter “Index”) is conceptually the number of 17-year-olds (children on the cusp of adulthood) who have never experienced the break-up of their biological parents. That is, the biological parents of these children have remained together (physically) to rear their children. The household, in this sense, is *intact*. The standard technical formulation of the Index is given in a section to follow entitled “Standard Formulation of the Index of Family Belonging and Rejection.”

That technically formulated Index (hereafter, “uncorrected Index”) contains systematic errors (leading to inaccuracies), largely because the survey on which it relies, the American Community Survey (ACS), only tracks the *biological relationship* of children to (at most) *one* of their parents. That relationship is to a parent identified as the head of household (or, “the householder”).

The uncorrected Index must thus look to the years of marriage of a couple and compare the time of marriage with the birth year of their child. Comparing this date with that of the child’s birth allows for an Index to be derived (whether the child of one biological parent was in an always-intact married family). ACS adequately represents the year of latest marriage of a couple.

Merely comparing the *year of marriage* to the child’s age is an inaccurate way of identifying a child’s *biological* parents. Biological parents of children may marry many years after the child’s birth. Additionally, even within just a couple years of a child’s

---

<sup>1</sup> Henry Potrykus, Ph.D., Senior Fellow, MARRI. Correspondence may be addressed to the author at [hgp@frc.org](mailto:hgp@frc.org).

birth, a mother may marry a man who *is not* the biological father of the child. This latter case may be something like an early childhood adoption. It regardless departs from the standard formulation of the Index.

In all of what follows, however, we still eschew the potential case of quick *re*-marriages after a child’s birth. These “double-shot” marriages (if weddings near a child’s birth be termed “shotgun marriages”), because they are of such quick succession, confound the ability to tie marriage to biology. In fact, the corrected Index (introduced below) also cannot track the biology of previous marriages.<sup>2</sup>

An additional technical point in the formulation of the uncorrected Index—that the age of the child is not tightly restricted to 17 years of age—has an additional effect on the measurement of the Index.

For each year that teenagers age, somewhat more than 1 percent experience the break-up of their intact, biological family (for teenaged children of married couples in their mid-teens, around 1.5 percent per year experience a divorce).

These occurrences effect a cancellation of errors: The uncorrected Index assumes too few biological parents eventually marry (it cannot track these couples over time). At the same time, the uncorrected Index *averages* the Index for children in the last three years of their childhood.<sup>3</sup> The younger ages have higher Index values: a higher level of intactness (sampling 15- and 16-year-old children, and not only 17-year-olds) *partially* cancels out *not* measuring biological marriages that take place later than two years after the birth of a child (measuring too low a value of the Index). This partial cancellation is investigated quantitatively in the present empirical analysis.

## Considerations on Correcting the Index Bias

Needless to say, with these varied family events influencing a *measured* Index (later biological marriage, early non-biological marriage, even separation and successive *re*-marriage back into biological families), a precise and accurate value of the Index is difficult to obtain. This work, with the aid of a survey which tracks biological relationships to both parents, quantifies the largest systematic errors in the uncorrected Index. The survey it exploits is smaller in size, so its findings will be less precise than those of the uncorrected Index, which employs the (massive) ACS.<sup>4</sup>

---

<sup>2</sup> Achieving this would require tracking a *history* of biological *relationships*, instead of simply marriage dates. Neither survey discussed here attempts this.

<sup>3</sup> For some analytic purposes, the sample size may be too small if the uncorrected Index only sampled children on the cusp of adulthood: 15- to 17-year-olds are sampled, as opposed to only 17-year-olds.

<sup>4</sup> Whereas ACS allows for a measured Index with tighter than three tenths of a percentage point precision, SIPP (introduced below) allows for measured Index values with the precision of about a percentage point. SIPP itself is a very large survey (measuring about 25,000 children per survey): This lack of precision is more of an indication of the rigors required in measuring such a childhood Index.

Even so, not every marriage-like relationship experienced by the biological parents of a child is tracked. Thus, as mentioned above, the children of re-acquainted biological parents count as *part* of the corrected Index of Family Belonging and Rejection. This re-acquaintance can occur after either separation or divorce.

Last, an empirical (statistically rigorous) attempt is made to quantify the risk of divorce in the final years of childhood (as already reported above). This quantification allows for the most accurate *homing in* on the number of children on the *cusp of adulthood* who have never experienced the break-up of their biological parents.<sup>5</sup>

## Standard Formulation of the Index

The designers of the uncorrected Index describe its formulation as follows:

“The procedure used to estimate the percentage of American adolescents aged 15 to 17 living with both of their married biological parents from the 2008 - 2011 American Community Survey Public Use Microdata Sets began by locating all persons in the public use data file who were in the target age range. We then checked the relationship of the teenager to the reference person of the household. (The reference person was the adult in the household in whose name the house or apartment was owned or rented.) If the teenager was coded as the biological son or daughter of the reference person, we checked to see if the parent was coded as being currently married. If so, we checked the date of the parent's most recent marriage. Was the marriage date before the year of the teenager's birth, or within two years of the birth year? If so, he or she was deemed to be living with both parents, who were continuously married throughout the teenager's childhood.

“If the teenager was described as the grandchild of the reference person, we checked to see if he or she was coded as ‘child in married-couple subfamily.’ If so, the teenager was deemed to be living with both married parents in a multigenerational family. We followed a similar procedure if the adolescent was described as the brother or sister or ‘other relative’ of the reference person, or as a roomer or boarder, housemate or roommate, or ‘other non-relative.’ So long as the teenager was also coded as ‘child in married-couple subfamily,’ he or she was deemed to be living with both married parents.

“Teenagers who were the biological child of the reference person but whose parent was divorced, separated, or never-married were classified as not living with both married parents. Likewise, if the teenager's birth antedated the year of the reference person's latest marriage by more than two years, the teenager was classified as not living with both parents but, rather, in a bioparent-stepparent family. If the parents were not married but cohabiting, the teenager was classified as not living with both married parents.

---

<sup>5</sup> Biological intactness is not the equivalent of intact marriage. Intact, cohabiting families are categorically different from intact, married families. However, these are also categorically different from families that do not enjoy biological intactness and differ from family structures that are traditionally considered “broken.”

“Teenagers who were described as the adopted son or daughter, stepson or stepdaughter, or foster son or foster daughter of the reference person were classified as not living with both married parents. Adolescents living in group quarters (e.g., correctional institution, halfway house) were classified as not living with both married parents. The number of teenagers living with both married bio-parents was divided by the total number of adolescents aged 15 to 17 in order to derive the percentage living with both parents.

“This rather complicated procedure is necessary because the 2011 ACS questionnaire only asks about a teenager's detailed relationship to the reference person, and not to the reference person's spouse or partner. Thus, we must infer that relationship by looking at the reference person's marital history information.”

The designers of the uncorrected Index go on to state: “We know this procedure is not 100 percent accurate. It may be, for example, that even though the parents were married throughout the teenager's childhood, one of the partners in the marriage is not, in fact, the biological parent of the teenager. Nonetheless, it is the best national data on the history of the families of the United States.”

It is the purpose of this study to analytically characterize this “gap” in knowledge.

## **Survey of Income & Program Participation Review**

As alluded to above, the survey used in this investigation, the Survey of Income and Program Participation (SIPP), tracks the *biological relationship* of children to their parents, *if the parents are present* in the sample household. This biological relationship forms the basis of the so-called corrected Index. In addition to knowing about present biological relationships, dates of marriages and divorces are tabulated for the household members. These dates constitute a marriage history in the SIPP.

This marriage history allows for an interesting profile to be developed on the age at which marriage and divorce are experienced *by children born out of wedlock*. Biological eventual marriage (that is, marriage that eventually occurs – after the child's birth) may be so tracked. Eventual marriage in general (biological or not) may likewise be tracked. Age at parents' divorce (in general) may be so tracked. Age at divorce of *biological* parents may not be tracked. One does not know the biological relation of any parent that was divorced from the family and moves away. This loss of information on a biological parent continues to hold, obviously, upon any remarriage to a stepparent.

With this information, let us summarize what of the Index, and the biasing terms, may be deduced:

The biological Index of Belonging (a corrected Index) itself is directly estimable.

This correction to the uncorrected Index estimates 1) the number of biological parents who eventually marry and 2) the number of those biological parents who eventually divorce:

- 1) Those biological parents who eventually marry and stay married are directly estimable.
- 2) Those biological parents who eventually divorce may only be estimated by the fraction of *all* parents that eventually divorce relative to all parents that eventually marry (a cumulative divorce rate) times the fraction of those *biological* parents that eventually marry and stay married relative to *all* those that eventually marry and stay married, regardless of biological affiliation (i.e., what fraction of all marriages of parents whose biological nature *may* be traced [irrespective of whether they are biological or not] *are* marriages of biological parents). This computation is necessary because the divorces of biological parents (only) cannot be segregated: the SIPP does not code (biological relationship) information on these broken relationships.

Combining the uncorrected Index (Index lacking biological information on parents that marry more than two years after a child’s birth) with the number of biological parents who eventually marry and stay married gives a corrected Index: Term (1) constitutes the correction to the biased Index.

Interestingly, the Index derived from the SIPP does not require that the parents eventually (ever) marry. It can, in principle, track (retrospectively—though there is a longitudinal component to the SIPP) long-term intact cohabitating relationships—so long as there is a child present. One might term these relationships “common-law marriages.”

These longer-term cohabitations are important to the correction of the Index. As will be seen in the profiles displayed below, they are in some degree quantifiable in terms of their transition to marriage. Tracking their transition to singlehood (usually the transition of the mother to that state) is beyond the scope of the present analysis.

## Results

**The Corrected Index.** From the SIPP (2008), the corrected Index for 17-year-olds is estimated to be around 47 to 49 percent (for non-Hispanic whites, it is around 53 to 55 percent). This is nearer the uncorrected estimate than the pure unaccounted-for marriages would suggest (see below) because of the cancellation of errors (later unaccounted-for biological marriages [which *do not* experience a divorce] cancelling against non-biological marriages early in the child’s life incorrectly counted as biological marriages and also cancelling against unaccounted-for divorces that will happen in a child’s late teens [ages 15 and 16]).

**Estimating the Index Bias.** Next, from the SIPP, 1)  $34 \pm 2$  percent of children born more than two years out of wedlock see their biological parents eventually marry

and stay married (the percentage for those born any time before wedlock is  $39 \pm 1.5$  percent).

From the next four statistics on children born more than two years out of wedlock, one can compute the bias term (2) above:

From the SIPP, the percentage of children born more than two years out of wedlock that eventually see their biological parents marry and not divorce is estimated to be  $37 \pm 2.5$  percent. (The percentage for those born any time before wedlock is  $44 \pm 2$  percent.)

From the SIPP, the percentage of children born more than two years out of wedlock that eventually see *any* parents marry and not divorce (irrespective of biological affiliation) is estimated to be  $60 \pm 3$  percent (the percentage for those born any time before wedlock is  $64 \pm 2.5$  percent).

From the SIPP, the percentage of children born more than two years out of wedlock that eventually live in a married household is estimated to be  $69 \pm 2$  percent (the percentage for those born any time before wedlock is  $75 \pm 2$  percent).

From the SIPP, the percentage of children born more than two years out of wedlock that eventually experience divorce (of the first household into which their parent marries) is estimated to be  $15 \pm 2$  percent (the percentage for those born any time before wedlock is  $20 \pm 2$  percent).

Of the last five statistics, combine the first four according to (2): 21 percent of children born more than two years out of wedlock whose biological parents eventually marry eventually see their parents divorce, if the divorce rate estimated by (2) holds.<sup>6</sup>

21 percent (2) of 43 percent (9 percent) leaves 34 percent (1, above). This implies that the biological parents of 43 percent of children who were born more than two years out of wedlock *were married at some point in time*.

According to the SIPP, about 29 percent of all children are born more than two years out of wedlock (34 percent of children are born at least a year before wedlock). Combined, 34 percent (1) of 29 percent (the relevant population potentially biasing the uncorrected Index) gives that the bias on the Index is 10 percent.

**Decomposition of Index Bias into Components.** The uncorrected Index *quoted* (quoted as 46 percent in previous reports) is 1 to 3 percentage points off of the corrected biological Index ( $48 \pm 1$  percent, above). However, this difference does not consider that, of children that are born more at least a year out of wedlock, perhaps 20 percent of those who belong *biologically* nonetheless never see their parents marry (while

---

<sup>6</sup> There are plausible reasons it may not.

they are still children). 20/80 of 10 percent [the *odds* of being a biologically belonging child of parents who never marry] is more than 2 percent. This is a reasonable *bound* on the systematic difference between the biological corrected Index [no marriage needed] and the bias corrected Index [by 1, above]. This number is highly uncertain (20 percent in fact is 1 standard deviation of the estimator - see the Technical Appendix).

With this consideration in mind, the quantified bias ( $10 - 2 = 8$  percent) allows in turn a rough quantification of the other terms influencing the uncorrected Index. First, younger ages have higher Index values: a higher level of intactness—sampling younger children—biases the Index up by 1.5 percent or so. Thus, early non-biological marriage (and any even separation and successive *re*-marriage back into biological families; note the foregoing paragraph) may have biased the Index in the upward direction by something approaching 3.5 percentage points ( $8 - 1.5 - 3$  [from the beginning of the previous paragraph] = 3.5).

Around 19 percent of unwed recent mothers marry. Around 12 percent of unwed recent mothers marry the biological fathers of their children. (Above 60 percent of all marriages of unwed recent mothers are to the biological fathers of their children. All of this is derivable from the SIPP.) 3.5 percent (from the previous paragraph's deductions) is in line with those statistics, if something like half of these mothers divorce their children's fathers ( $19$  percent -  $12$  percent =  $7$  percent;  $7$  percent divided by  $2$  is  $3.5$  percent). As A) these numbers are commensurate, B) somewhat less than half of these "shotgun" marriages end in divorce, and C) given the bound on the uncertainty in biologically belonging children who never see their parents marry (2 percent, two paragraphs above), the corrected Index may be most likely (in the statistical sense) chosen to the higher end of its range (that is, about 49 percent).<sup>7</sup>

---

<sup>7</sup> The 3.5 percent residual is too small, because less than half of the 7 percent's marriages fail. Also, there is uncertainty concerning biologically belonging children who never see their parents marry. Hence, the original bias (10 percent) should be larger [after taking out all other systematic errors, i.e., in the  $10 - 2 - 1.5 - 3 = 3.5$  relationship]. Consequently, the corrected Index itself may be most likely chosen to the higher end of its range (48 to 49 percent, as opposed to 47 to 49 percent, because the corrected Index is 'pushed' higher above the uncorrected Index by the original bias being reckoned too small: The uncorrected Index, while inaccurate, is very precise, and so is a good baseline off of which to deduce a bias-corrected Index). Of course, all these computations are subject to error, of around a percentage point or so itself.

## Technical Appendix: Deduction of Eventual Marriage & Divorce Fractions

We determine right-censored Kaplan-Meier survival rates for sub-populations of children of interest. The censoring event is the survey itself: the child measured at the time of the survey may not have reached 17 years of age. See the text for the sub-populations whose hazards are tracked. Plainly, this is not a true cohort survival rate, in the sense of a pure cohort of children aging over a tracked period. It is a “total survival rate” for a population in a given year, analogous to the “total fertility rate” of a population in a given year.

SIPP is not an unbiased sample of the U.S. population, so we apply the SIPP-provided weights to the Kaplan-Meier statistic. The standard errors are quoted as twice the unweighted sample-size Greenwood statistics (no clustering has been applied). The standard errors most usually quoted are Greenwood standard errors of the age 17 survival rates.

In the deductions from these elemental survival rate calculations (primarily (2)), one may perform simple interval arithmetic to give a standard error for the composite.

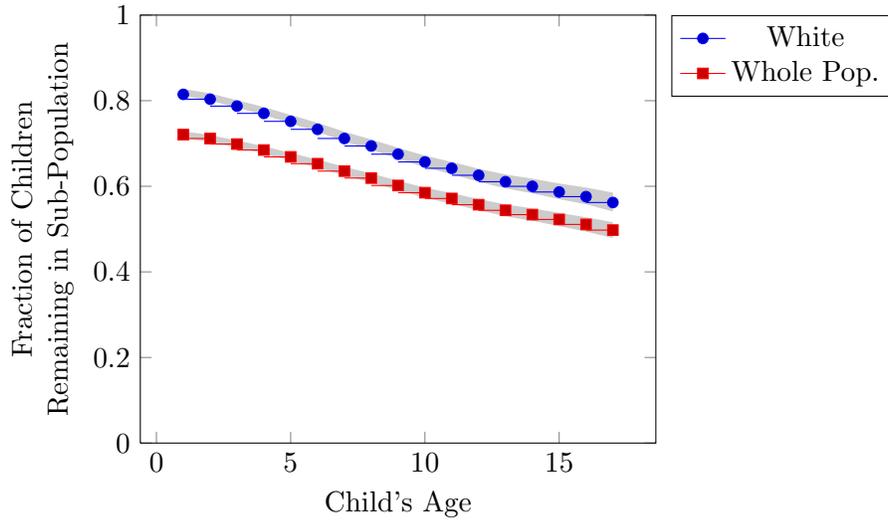
What follows are the graphs of the cumulative survival rates, in the order that the populations are introduced in the text. Note that we chart the survival rates of these sub-populations across all years after the out-of-wedlock child is born, not only those years relevant to the bias correction (years greater than 2 beyond the child’s birth). These plotted survival rates are more physically comprehensible.

---

**Chart 1** Survival of Child's Intact Biological Family

1- to 17-year-olds, SIPP 2008

---

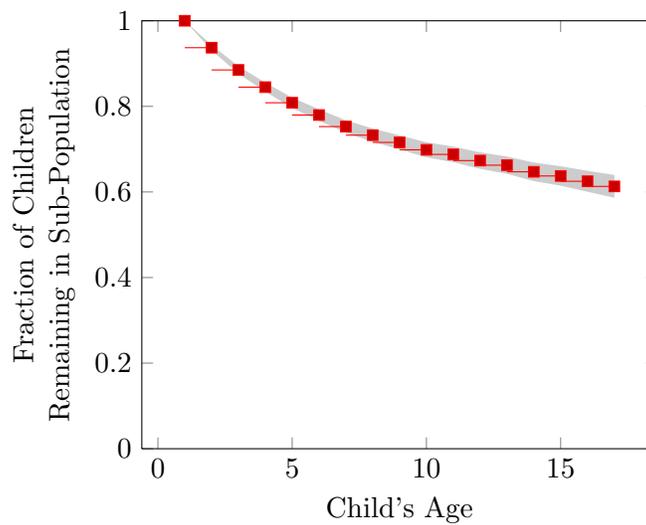


---

**Chart 2** Remainder of Child's Biological Parents *Not* Marrying

1- to 17-year-olds born out of wedlock, SIPP 2008.

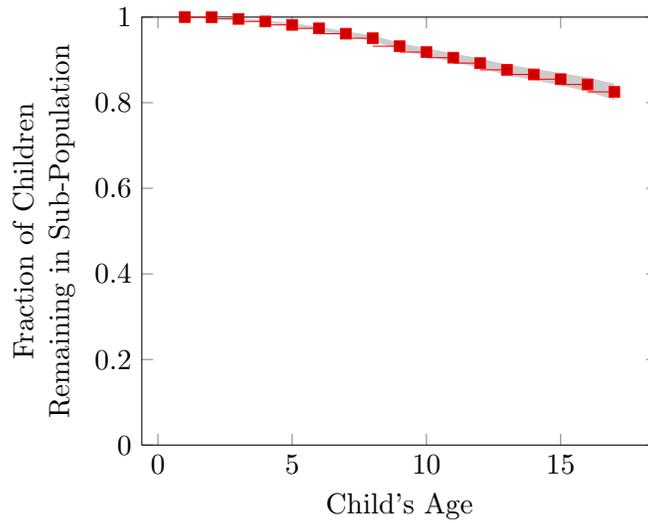
---



---

**Chart 3** Remainder of Children *Not* Experiencing a Divorce  
1- to 17-year-olds born out of wedlock, SIPP 2008.

---



---

**Chart 4** Remainder of Children *Not* Experiencing a Marriage  
1- to 17-year-olds born out of wedlock, SIPP 2008.

---

