

## Table 2: Baseline Cross-Sectional Estimates by College

This table reports our baseline estimates of parents' and children's income distributions by college. We calculate college-level values as means over students in the 1980, 1981 and 1982 birth cohorts. When data for a college from any of these cohorts are incomplete, we use data from 1983 and 1984 cohorts to obtain an estimate. See the codebook for Online Data Table 1 for definitions of income variables, college attendance, and other details.

In some cases, multiple colleges share a single identifier in the tax records (e.g., the campuses of the University of Massachusetts or University of Illinois systems). In these cases (denoted by multi = 1 in the dataset), we report averages over all students attending the relevant group of colleges. For these groups of colleges, we assign locations based on the location of the largest school/campus with the group of colleges.

Variable	Description
super_opeid	Institution OPEID / Cluster ID when combining multiple OPEIDs: -99 = late goers (attended college between 23-28) -9 = Never attended college (up to and including year 2013) -1 = attending a college with insufficient data (less than 100 students on average or incomplete data) >0 = specific colleges
name	Name of college (or college group)
type	Type : 1 = public 2 = private non-profit 3 = for-profit
tier	Selectivity and type combination (see Table 6 for more detailed descriptions of these groups): 1 = Ivy Plus 2 = Other elite schools (public and private) 3 = Highly selective public 4 = Highly selective private 5 = Selective public 6 = Selective private 7 = Nonselective 4-year public 8 = Nonselective 4-year private not-for-profit 9 = Two-year (public and private not-for-profit) 10 = Four-year for-profit 11 = Two-year for-profit 12 = Less than two year schools of any type 13 = Attending college with insufficient data 14 = Not in college between the ages of 19-22
tier_name	Name of college tier

iclevel	Four-year or two-year college 1 = Four-year 2 = Two-year 3 = Less than Two-year
region	Census region: 1 = Northeast 2 = Midwest 3 = South 4 = West
state	State
cz	Commuting zone ID
czname	Commuting zone name
cfips	Combined state and county fips code
county	County
multi	Indicator that equals 1 if multiple colleges (IPEDS Unit IDs) are grouped in this Super OPEID
count	Average number of kids per cohort
female	Fraction female among kids
k_married	Fraction of kids married in 2014
mr_kq5_pq1	Mobility rate (joint probability of parents in bottom quintile and child in top quintile of the income distribution)
mr_ktop1_pq1	Upper-tail mobility rate (joint probability of parents in bottom quintile and child in top 1% of the income distribution)
par_mean	Mean parental income
par_median	Median parent household income (rounded to nearest \$100)
par_rank	Mean parental income rank
par_q[PARQUINT]	Fraction of parents in an income quintile [PARQUINT]. 1 is the bottom quintile and 5 is the top.
par_top[PCTILE]pc	Fraction of parents in the top percentile [PCTILE]. For instance, par_toppt1pc refers to parents in the top 0.1% of the income distribution.
k_rank	Mean kid earnings rank
k_mean	Mean kid earnings
k_median	Median child individual earnings in 2014 (rounded to the nearest \$100)
k_median_nozero	Median child individual earnings among positive earners in 2014 (rounded to the nearest \$100)
k_0inc	Fraction of kids with zero labor earnings
k_q[KIDQUINT]	Fraction of kids in an income quintile [KIDQUINT]. 1 is the bottom quintile and 5 is the top.
k_top[PCTILE]pc	Fraction of kids in the top percentile [PCTILE]. For instance, top1pc refers to children in the top 1% of the income distribution.

k_rank_cond_parq[PARQUINT]	Mean kid earnings rank conditional on parent in quintile [PARQUINT]
kq[KIDQUINT]_cond_parq[PARQUINT]	Probability of kid in quintile [KIDQUINT] conditional on parent in quintile [PARQUINT]
ktop1pc_cond_parq[PARQUINT]	Probability of kid in top 1% conditional on parent in quintile [PARQUINT]
k_married_cond_parq[PARQUINT]	Fraction of kids married conditional on parent in quintile [PARQUINT]
shareimputed	Share of count-weighted data that was imputed using information from the 1983-84 cohorts
imputed	Indicator if any data for that super_opeid was imputed