Mobility Report Card: Angelina College

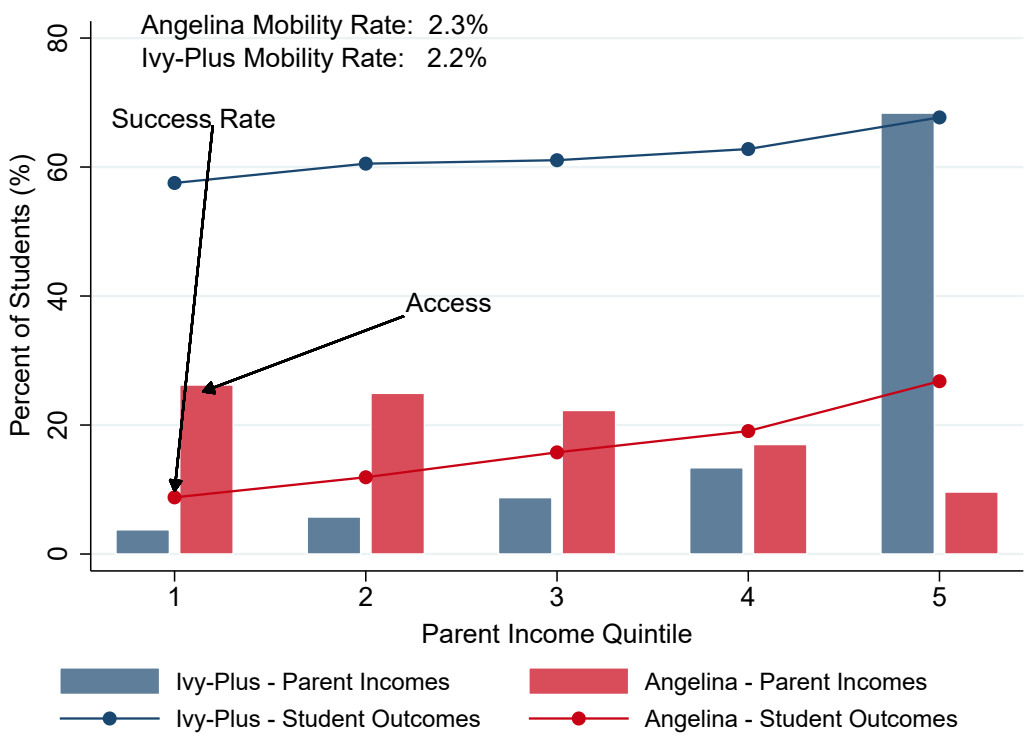
Raj Chetty, John N. Friedman, Emmanuel Saez, Nicholas Turner, Danny Yagan

Which colleges in America contribute the most to helping children climb the income ladder? How can we increase access to such colleges for children from low income families? We take a step toward answering these questions by constructing publicly available *mobility report cards*: statistics on students' earnings in their early thirties and their parents' incomes for each college. We estimate these statistics using de-identified data from the federal government covering all students attending college from 1999-2013, building on the Dept. of Education's College Scorecard. This report presents statistics from the Mobility Report Card for Angelina and selected groups of colleges, including the Ivy-Plus (Ivy League plus MIT, Stanford, Chicago, and Duke) and the 270 four-year colleges with six-year graduation rates consistently above 70% ("Top 270").

Parents' Incomes and Students' Earnings Outcomes

Figure 1 below presents statistics for the students born in 1980-82. The bars show the fraction of parents who come from each quintile of the income distribution; we label the fraction of parents from the bottom 20% as "access." The lines show the fraction of students from each of those quintiles who make it into the top fifth as adults, which we label as "success rates." The product of "access" and the "success rate" for children from families in the bottom quintile is the "mobility rate": the fraction of students who come from families in the bottom income quintile and end up with earnings in the top quintile.

Figure 1: Mobility Report Card, 1980-1982 Birth Cohorts



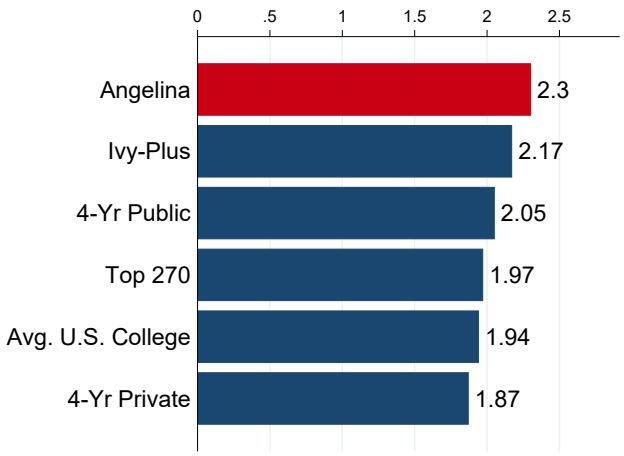
College Mobility Rates

Figure 2(a) shows the mobility rate for Angelina compared with average mobility rates across Ivy-Plus schools, Top 270 schools, all four-year public schools, all four-year private non-profit schools and all schools in our study. In Figure 2(a), where success is measured as reaching the top 20% of the income distribution, Angelina ranks in the 78 percentile nationally. In Figure 2(b), where success is measured as reaching the top 1% of the income distribution, Angelina ranks in the 90 percentile nationally.

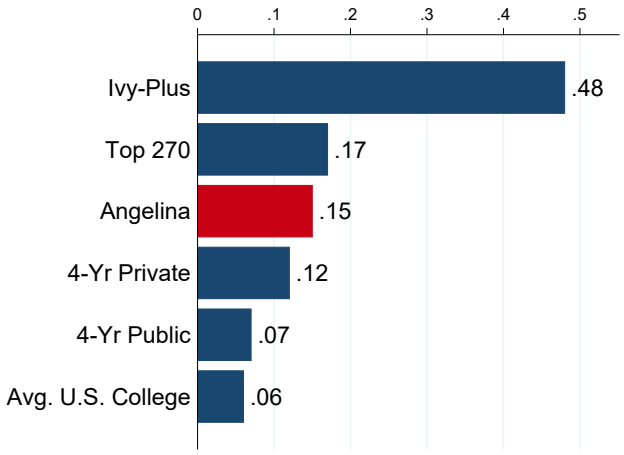
$$Mobility\ Rate = Success\ Rate \times Access$$

Figure 2: Mobility Rates, 1980-1982 Birth Cohorts

(a) Top 20%



(b) Top 1%

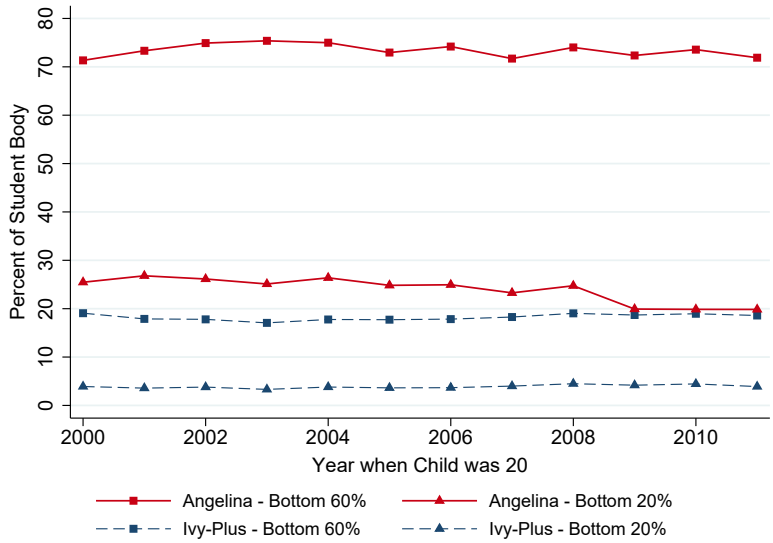


Changes over Time

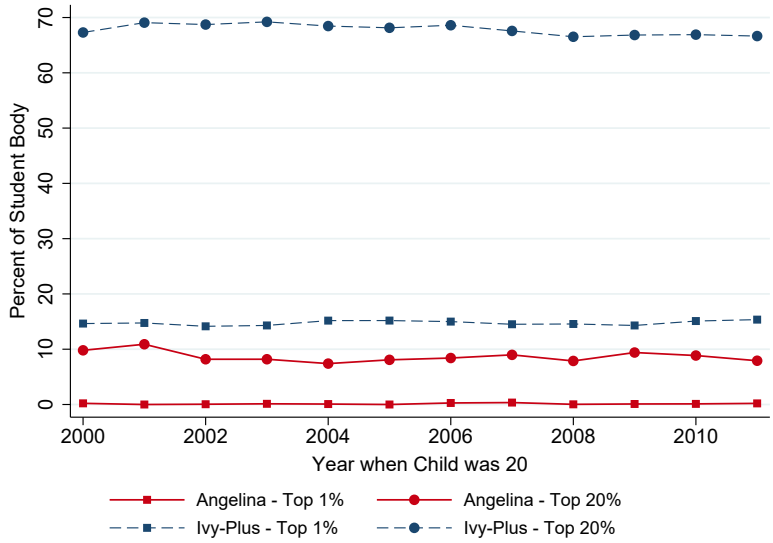
The Mobility Report Card also includes statistics on how the parent income distribution at Angelina has evolved since 2000, when our data begin. Figure 3(a) shows the share of students from the bottom 20% and bottom 60% of the income distribution, while Figure 3(b) shows the share of students from the top 1% and top 20%.

Figure 3: Trends in Parent Household Income, 1980-1991 Birth Cohorts

(a) Lower Income Shares



(b) Top Income Shares



Appendix: Sources, Definitions and Methodology

A complete description of our data sources and methodology is given in our academic paper, which is available online here. The full public data are available here.

Data. Our analysis builds upon the datasets used to construct the Department of Education's College Scorecard. We use data from de-identified federal income tax returns (1996-2014), including 1098-T forms for college attendance, and the National Student Loan Data System from the Department of Education (1999-2013) to obtain information on college attendance, students' earnings in early adulthood, and their parents' incomes.

College Definitions. We say a student attended a college if it is the institution she attended most frequently between the ages of 18 and 22 (from the year in which she turns 19 to the year in which she turns 22). If a student attends multiple colleges for the same number of years, we choose the college she attended first. Our results are not meaningfully different when we use other reasonable definitions of college attendance (for more on this see Section II.B of our paper).

Income Definitions. We define parents as the most recent tax household to claim a child during the years when she is aged 12-17. We then measure parent income as total pre-tax income at the household level. For most households, this is their adjusted gross income, as reported on their 1040 tax return. For non-filers, we impute based on information returns such as W2s. If a family has no tax or information returns of any kind, we record income as $0. We then average parents' family income over the five years when the child is aged 15-19 to smooth transitory fluctuations and obtain a measure of resources available at the time when most college attendance decisions are made. We then assign parental income percentiles by ranking each household based on mean income relative to all other households with children in the same birth cohort.

Our primary measure of student income is total pre-tax individual earnings in 2014, defined as the sum of wage earnings and net self-employment income as reported on Form 1040 or information returns. We assigned a percentile rank relative to the distribution of earnings for all other children in the same birth cohort nationally. For students in our core sample, who were born in the 1980-1982 cohorts, income is therefore measured in the early-to-mid thirties, by which time our data suggest that income percentile ranks have stabilized (see Figure II in our paper). For more information on income definitions, see Section II.C of our paper.

Imputation. Our primary statistics average across data for students born in the 1980-82 cohorts. If a college is missing one or more of these years of data–typically because of incomplete reporting of 1098-T forms–we impute values for the missing cohorts using data from the 1983-84 cohorts. We do this to increase the number of colleges for which we can report data. The main results of our paper hold if we exclude imputed data. For more information on the imputation procedure, see Section II.D of our paper.

Estimation. Following established disclosure standards for the use of tax data (such as the standard of aggregating over 10 or more tax units when disclosing statistics), we report estimates for each college that are based on tabulations that aggregate across several colleges. The estimation error is typically smaller than the year-to-year fluctuations in the underlying statistics. For more information on our estimation procedure, see Appendix C of our paper.