# The Equitable Value Explorer – Data Documentation

Note: This data documentation reflects data updates as of May 2023.

## Contents

Overview	1
Universe	2
Missing Data	2
Post-College Earnings Outcomes	2
Economic Value Thresholds	3
Total Student Investment	4
Economic Value Index (EVI)	5
Equitable Value Contribution (EVC)	6
Other Contextual Variables	6

## Overview

The publicly available institution-level data underlying the Equitable Value Explorer are assembled from a variety of sources, including the Integrated Postsecondary Education Data System (IPEDS), the U.S. Census Bureau's American Community Survey (ACS), and College Scorecard data. This documentation describes the dataset construction, identifies the sources of each variable, and explains the methodology behind constructed variables. Where applicable, this document also includes subsections for relevant variables describing how more robust student-level data, such as those used by the University of Texas (UT) System, can produce more precise results.

The Equitable Value Explorer is meant to provide information to postsecondary leaders and stakeholders using institution-level data to inform assessment of the economic value provided to students by postsecondary institutions. More importantly, it shows how equitably this value is delivered across different subgroups of the student body. New metrics constructed to measure economic value are listed and described alongside other contextual variables included in the Explorer. This guide includes the underlying data source(s) and cohort(s) for each metric. Users can refer to the relevant College Scorecard and IPEDS documentation files (listed below) for detailed descriptions of the underlying variables.

Dataset	Document Type	Link
College Scorecard	Data Dictionary	Link
	Technical Documentation	<u>Link</u>
IPEDS	Overview	<u>Link</u>
	Glossary	<u>Link</u>
American Community Survey (ACS), downloaded through IPUMS USA	Technical Documentation from ACS	<u>Link</u>
	IPUMS USA Overview	Link

## Universe

The Equitable Value Explorer's data universe is comprised of data from postsecondary institutions present in the 2021-22 College Scorecard institution-level data file (MERGED2021\_22\_PP), published in April 2023. This file includes all IPEDS institutions that participate in Title IV aid programs, excluding administrative offices, schools that have missing data on degree and certificate completions (preventing institution category classification), or those that have no valid Office of Postsecondary Education identification number (OPEID). This universe was pared down by restricting it to the subset of institutions a) whose predominant credential awarded is an undergraduate credential, b) who are located in the 50 states or the District of Columbia, and c) who report positive undergraduate enrollment. Altogether, 5,615 institutions are in the data universe.

### **Missing Data**

As explained in detail below, the constructed economic value metrics measure the earnings outcomes of students receiving federal financial aid—published through the College Scorecard—relative to median earnings of a comparison population drawn from ACS. Value metrics also include estimates of the costs students incur to earn a credential at each institution. Earnings data in the Scorecard are based on a match of all federally aided students to yearly administrative earnings data at the Social Security Administration. To protect privacy, data are suppressed when the number of students in a cohort is deemed too small.<sup>1</sup>

Data on costs are estimated from IPEDS, based on data on reported charges, student living arrangements, and graduation rates. In estimating students' total investment in college, we only included colleges operating on a standard academic calendar, which reports cost of attendance to IPEDS across a full academic year. Some institutions operate on a calendar year that differs by program or offer programs on a continuous basis, and report costs to IPEDS for their largest programs. Due to the complexity of integrating program-level cost data, which cover differing lengths of time, we do not include cost information or calculate Threshold 0 for these institutions.

Of the 5,615 total institutions in our universe, 4,518 institutions have valid earnings data and 2,921 have sufficient data to enable the construction of Threshold 0 (see section on "Economic Value Thresholds" below).

## Post-College Earnings Outcomes

The Postsecondary Value Framework is designed to measure the share of an institution's students overall and by different demographic subgroups—whose earnings exceed specific thresholds. Because the College Scorecard does not provide data on the percentage of students earning above these thresholds, the Equitable Value Explorer compares the median overall earnings of students against the thresholds (see threshold discussion below). The measures of overall earnings used in the tool are based on median earnings measured 10 years after students enter an institution, from the College Scorecard. These earnings data shown were collected in calendar years 2019 and 2020 for students who first enrolled between 2008-09 and 2009-10.

The Scorecard also contains median earnings separately for select student subgroups: data are available separately for men and women, and for three different family income groups for dependent students

<sup>&</sup>lt;sup>1</sup> While the precise number of students required for privacy suppression can vary across metrics due to a number of methodological factors, the Department of Education (ED) reports that public-facing elements in the Scorecard Tool suppress data for institutions with fewer than 30 students in the denominator to ensure representative data.

only. College Scorecard earnings data are not currently disaggregated by race/ethnicity, though this will be possible in the future with the addition of race/ethnicity questions to the FAFSA. All variables reported in dollars are adjusted to real 2022 dollars using the Consumer Price Index for All Urban Consumers (CPI-U).

The UT System can disaggregate median earnings outcomes by race, income, and gender, as sample sizes allow, and can calculate the percentage of students earning above a given threshold. UT System data can also be calculated at the institution or program level. Because the system has historical data on student completion, it also can define earnings cohorts based on exiting students rather than entering students and can disaggregate earnings outcomes for program completers and non-completers, overall as well as by program of study. The UT System can measure earnings outcomes up to 15 years post-exit. To achieve maximum coverage of outcomes and disaggregates across programs, the system combined 16 cohorts of data.

## Economic Value Thresholds

The Equitable Value Explorer measures institution-level earnings (and, in the case of the UT System tool, institution- and program-level earnings) against four economic value thresholds, referred to as T0, T1, T2, and T3, defined in Table 1 below. The thresholds are estimated using data from ACS, based on a combined 5-year file that aggregates data from 2017 to 2021. These threshold earnings estimates are based on individuals with positive earnings between the ages of 22 and 40<sup>2</sup> and not enrolled in a postsecondary institution in the three months prior to the survey interview. The ACS earnings measure INCEARN includes income from wages and from any self-owned business or farm. While business and farm earnings can include negative figures, we exclude any zero or negative incomes from our threshold calculations. These earnings values are reported in 2021 dollars in the ACS and adjusted to 2022 dollars using the annual CPI-U.

Threshold	Measurement
T0 "Minimum Economic	Median earnings of workers with positive earnings in the same state
Return"	the institution is located and with a High School Degree (or GED) as
	their highest level of educational attainment, plus yearly amortized
	total cost of obtaining the credential.
T1 "Earnings Premium"	Median earnings of workers with positive earnings in the same state
	with the same level of degree (BA or AA) as the predominant
	credential awarded by the institution. For institutions that award less-
	than 2-year credentials, the earnings threshold is set at 0.89 times the
	threshold for Associate's degree holders. <sup>3</sup> For University of Texas
	System data, the median earnings are calculated at the program level.

#### Table 1: Value Threshold Definitions

<sup>&</sup>lt;sup>2</sup> Data from the UT System as well as from the City University of New York (CUNY) system illustrate that income trends are similar whether younger or older age bands are used. An analysis of National Postsecondary Student Aid Study (NPSAS) data illustrates that the average age of completion is 27. As a result, the age range of 22-40 allows for measuring earnings from a broad range of students including those who complete a four-year degree within four years directly after high school, as well as non-traditional students who complete degrees later.
<sup>3</sup> The ACS does not include data on the earnings of certificate-holders, so this framework estimates certificate earnings. This factor—0.89—is the approximate proportional difference in earnings between those with certificates vs. Associate's degrees, measured in the Survey of Income and Program Participation by the Georgetown Center for Education and the Workforce.

T2 "Earnings Parity"	Median earnings of male (when measuring female earnings) or White (when measuring non-White earnings) workers with positive earnings in the same state with the same degree as the predominant credential awarded by the institution. (Data are unavailable separately for
	workers who were high-income in childhood). For institutions that award less-than 2-year credentials, the earnings threshold is set at 0.89 times the threshold for Associate's degree holders. <sup>4</sup>
T3 "Earnings Mobility"	The 60 <sup>th</sup> percentile of earnings among workers with positive earnings in the same state the institution is located.

Where sample sizes are sufficient, disaggregated thresholds also are calculated separately for specific subgroup populations, including gender and six race/ethnicity groups (White non-Hispanic, Black non-Hispanic, Asian and Pacific Islander, Indigenous, and Other). Thresholds cannot be calculated separately for individuals from low- or high-income backgrounds because the ACS does not collect data on individuals' childhood economic circumstances; longitudinal data are needed to calculate thresholds by income status.

## Total Student Investment

To estimate students' total investment in college, we use publicly available data from the U.S. Department of Education's IPEDS surveys. First, for each college, we calculate the net price paid each year by the typical student. Net price is defined as the total cost of attendance (including tuition and nontuition costs) minus grant and scholarship aid. This annual net price estimate is then multiplied by an estimate of the number of years it takes the average graduate to finish their credential, and the resulting measure of cumulative net price is amortized over a ten-year period to account for the cost of student loan interest.

The cost of attendance data in the Equitable Value Explorer are taken from the 2021 Institutional Characteristics IPEDS survey and include colleges' published costs for tuition and fees, books and supplies, room and board, and other expenses for first-time, full-time (FTFT) students. The latter two components are estimated separately for students living on-campus, off-campus with family, or off-campus without family, and we weight the average cost of attendance for an institution based on the distribution of its students' living arrangements. We also apply the estimated room and board expenses for students living off-campus, not with family to those living off-campus, with family, to acknowledge that many students living with family still purchase food and/or pay rent.<sup>5</sup> As discussed under "Missing Data" above, the data in the Equitable Value Explorer are limited to colleges that operate on a standard academic calendar and report cost of attendance to IPEDS across a full academic year.<sup>6</sup>

We estimate net price by subtracting average grant aid awarded to all FTFT degree-seeking students, as reported in the 2021 Student Financial Aid IPEDS survey, from the annual cost of attendance calculated above. Although IPEDS includes its own net price calculations, those estimates are limited to financial aid recipients and include incomplete costs for students who live with family.

Because the Threshold O calculation requires an estimate of students' total investment in college, we multiply the annual net price by the estimated number of years it takes the average graduate at a

<sup>&</sup>lt;sup>4</sup> See footnote 3.

<sup>&</sup>lt;sup>5</sup> IPEDS currently does not allow colleges to report room and board costs for students living with family.

<sup>&</sup>lt;sup>6</sup> Some institutions operate on a calendar year that differs by program or offer programs on a continuous basis, and report costs to IPEDS for their largest programs. Due to the complexity of integrating program-level cost data, which cover differing lengths of time, we did not include those colleges in our analysis.

particular institution to finish their credential. There are no institution-level data publicly available that track graduates' actual time to degree, so we estimate time-to-credential using data from the 2020 and 2021 IPEDS Graduation Rate (GR) surveys.<sup>7</sup> Estimated time-to-credential is calculated as a weighted average of the completion times among FTFT students graduating within 150 percent of the expected completion time. For predominantly bachelor's degree granting institutions, IPEDS collects the number of completers within 4, 5, and 6 years. We calculate estimated time-to-credential as the weighted average of the number of years it takes those students to complete, omitting both non-completers and those taking longer than six years to complete. At predominantly associate and certificate granting institutions, we estimate an analogous measure as the weighted average of the number of years to complete within 3 years.

For instance, a predominantly associate's-degree granting institution with 10 students who take two years to complete and 20 students who take three years to complete would have an estimated time-to-credential of 2.66 years.  $[(2 \times 10)+(3 \times 20)]/30 = 2.66$ .

In cases where colleges do not have available graduation rate data to estimate time-to-credential, we apply assumptions about expected time-to-credential. For institutions that report no completers within 150% of normal time to completion, we use the 150% timeframe (six years for four-year institutions, and three-years for two-year institutions) as their estimated time-to-credential. For institutions that do not report GR data to IPEDS, we assume that certificates take one year, and associate and bachelor's degrees take 125% of normal time to completion (2.5 years for AA; 5 years for BA).

Because approximately 70 percent of students who finish a bachelor's degree complete with student debt,<sup>8</sup> which incurs additional costs due to interest, we estimate the cost of student loan interest. We assume that cumulative net price is amortized over 10 years, using the 2021-22 federal undergraduate student loan interest rate (3.73%). Finally, this amortized amount is added to the median earnings of high school graduates from the ACS, resulting in a unique Threshold 0 value for each institution.

UT System data can produce a more robust estimate of the full cost of attendance. The data allow for calculating the actual net price paid by students in each term, accounting for fluctuations in tuition, grant aid, fees, and more. These more detailed data, in turn, allow for a more precise calculation of Threshold 0. This measurement is based upon data from four campuses in the UT System.

UT System data also includes actual, rather than estimated, time-to-credential for completers. The UT System can calculate the number of years a student is enrolled before leaving school, even if they do not complete a credential, providing a better baseline for understanding the typical financial amount a student without a credential will spend and need to recoup from employment. It also enables estimates of time-to credential/stop-out on a semester-by-semester basis, rather than an annual basis. These data are available at the institution and program level in the UT System.

## Economic Value Index (EVI)

The economic value index is designed to capture the fraction of a school's completers that attain at least a minimum economic return. It is the product of two components: the share of a subgroup's completers

<sup>&</sup>lt;sup>7</sup> We pool the 2020 and 2021 GR surveys in order to increase the number of completers in each category, which improves the reliability of estimates. This is particularly important for developing estimates of time to credential for race/ethnicity and gender subgroups of students.

<sup>&</sup>lt;sup>8</sup> Urban Institute. Understanding College Affordability: Borrowing. Retrieved from: <u>http://collegeaffordability.urban.org/covering-</u>expenses/borrowing/#/

whose earnings are greater than the T0 threshold described above, and the share of a school's completers represented by that subgroup. Without individual level earnings data, it is not possible to measure the share of students in any group earning above a threshold. Reflecting these data constraints, the Equitable Value Explorer suppresses the EVI of all institutions, showing only the completion element of the formula (which is available using IPEDS completion data).

The UT System data can, through their post-college earnings data, identify precisely which students have met each threshold and when. As a result, they can produce an accurate percentage of students meeting the thresholds, which also allows the production of the EVI. This is available at the institution and program level, and separately by race/ethnicity, gender, and income.

Equitable Value Index (EVI) = (% of student group completers passing Threshold 0) × (% of completers represented by selected student group)

## Equitable Value Contribution (EVC)

The equitable value contribution (EVC) is designed to measure the economic activity generated by postsecondary institutions whose students' earnings exceed the minimum return. EVC values are not currently included in the Equitable Value Explorer, except for the UT System (see below). A future update of this tool will include EVC values by gender and income but not race/ethnicity, because median earnings are not separately available by race/ethnicity in the College Scorecard.

With a robust Threshold 0 cost estimate and disaggregated post-college earnings estimates (for completers only), the University of Texas System data allows for precise estimates of the EVC. This is available at the institution and program level.

Equitable Value Contribution (EVC \$) = (actual median earnings of completers in student-group – Threshold 0) x (# of completers within student group)

We also report the percentage of the overall EVC generated by each student subgroup, which is determined by dividing the student subgroup EVC by the overall EVC:

Equitable Value Contribution (EVC %) = student-group EVC / Total Value Contribution<sup>i</sup>

### Other Contextual Variables

**Institutional sector** is based on the control of each institution (public, private nonprofit, and private forprofit) and the predominant degree awarded.

**Minority serving institution (MSI)** status is determined by the College Scorecard classifications, and includes institutions flagged as one or more of the following: Historically Black Colleges and Universities (HBCUs), Predominantly Black Institutions (PBIs), Alaska Native and Native Hawaiian serving institutions

(ANNHSIs), tribal colleges and universities (TCUs), Asian American Native American and Pacific Islander serving institutions (AANAPISIs), Hispanic serving institutions (HSIs), and Native American non-tribal institutions.

**Enrollment.** Full-year undergraduate unduplicated headcount enrollment is drawn from IPEDS 12-month Enrollment Survey and reflects enrollment at any point during the 2020-21 academic year. This data is the source used for all enrollment figures, including enrollment percentages for each race and gender group. Data on the proportion of Pell Grant recipients comes from the College Scorecard and reflects the percentage of all undergraduates who received Pell Grants in academic year 2020-21, as reported to IPEDS in 2021-22.

**Graduation Rate**. These data reflect the percentage of first-time, full-time students who receive a degree or credential within 150% of normal time to completion, disaggregated by race, gender, and Pell status. For four-year schools, the rates reflect pooled cohorts, from IPEDS data on completion outcomes for the 2014-15 and 2015-16 cohorts, reported to IPEDS in 2020 and 2021. For two-year schools, the cohort is based on pooled cohorts from 2017-18 and 2018-19.

**STEM.** Using the College Scorecard data on the share of awards in different Classification of Instruction Program (CIP) areas, % STEM combines the share of awards in each of the following fields: computer and information sciences, engineering, biological sciences, math, and physical sciences (pcip11 pcip14 pcip15 pcip26 pcip27 pcip29 pcip40 pcip41). CIP codes reported in the College Scorecard are based on data reported to IPEDS in 2021-22 and reflect student completions by field for academic year 2020-21.

**Cohort Default Rate.** The three-year cohort default rate from College Scorecard measures the share of borrowers entering repayment in a given cohort who default within three years. The CDR published in the College Scorecard reflects the default outcomes for borrowers who enter repayment in fiscal year 2019, as of 2021.

**Instructional Expenditures per FTE.** The reported per-full-time-student instructional expenditures, from College Scorecard. The calculation is based on IPEDS data collection year 2021-22, with expenditures reflecting FY2021 levels, and full-time equivalent enrollment for academic year 2020-21.

<sup>&</sup>lt;sup>i</sup> The Total Value Contribution represents the combined EVC's of *all* student groups, including non-target student populations.