Colorado Minimum Value Threshold - Work Update

Pursuant to HB22-1349 and the Colorado Commission on Higher Education (CCHE) Strategic Plan, CCHE has convened a technical working group ("Working Group") to:

1. "…develop student success measures that measure the progression of students through postsecondary education and the impact of postsecondary pathways on a student's career opportunities and success," and
2. Strategize to "create and maintain a statewide student success data system," to be administered by the Colorado Department of Higher Education (CDHE.)

The primary task of the Working Group has been to determine the most appropriate empirical model to estimate the minimum value threshold of postsecondary credentials in Colorado, and the best data elements to support this model. CCHE will approve the general framework and methodology for the minimum value threshold. In alignment with the goals outlined in the Strategic Plan, the output of this model will support **collaborative conversations** between CDHE, CCHE, and Colorado institutions of higher education to better understand outcomes for programs and identify potential actions that all stakeholders can take to increase student success.

The following overview provides a technical update on this work and a distillation of issues, conversations, and proposed solutions.
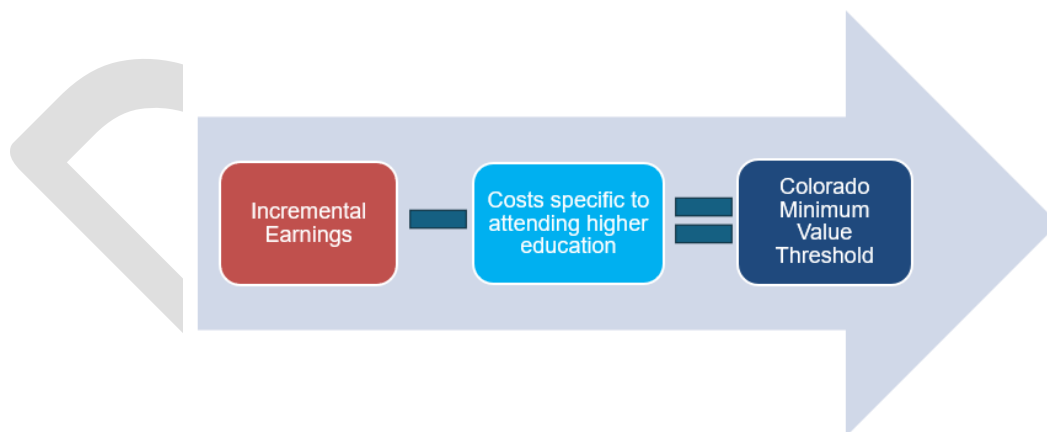
The Minimum Value Threshold (MVT) calculation will consist of three components:

1. **Realized earnings**, the earnings that students who participate in higher education realize after they separate from school,
2. **Counterfactual earnings**, the earnings that students who participate in higher education would have earned had they not attended school (i.e., if they had entered the labor force with only a high school degree.) These include lost labor market earnings while enrolled in school, and
3. **College costs**, the costs specific to attending higher education, including net tuition and required fees, books, and supplies.

MVT can be calculated at several different levels of aggregation: for the individual student, for the degree program, for the institution, for certain demographic groups, or for Colorado as a whole. The best strategy for calculating MVT at these different levels of aggregation is to calculate each of the three components for each individual student with available data, then aggregate those students into cohorts however we wish.

MVT can then be calculated at the cohort level, such as postsecondary **completers and non-completers**. We can take the central tendency (either average or median) of realized earnings for all students in the cohort, then subtract the central tendencies of counterfactual earnings and college costs for all students in the cohort. It is also possible to calculate other measures of economic value using this data, such as the percentage of students who outearn their counterfactual selves in any given year.

For the initial MVT calculation, we will use a combined cohort of students who first enrolled in higher education in 2007 or 2008. We will follow these students for fifteen years to assess their MVT from higher education. For the 2007 cohort, we follow the students through 2021; for the 2008 cohort, we follow the students through 2022. Moving forward, we will use trends and patterns established using these data to forecast the earnings and workforce outcomes of future cohorts of students, rather than rely solely on a backward-looking approach.

These cohorts will exclude any student who enrolled in graduate school during the measurement period, as well as any student who was enrolled in school for more than half of the measurement period. The cohorts will also include out-of-state students.

Proposed method of calculation

The following sections explain the proposed method of calculation for each of the components of MVT.
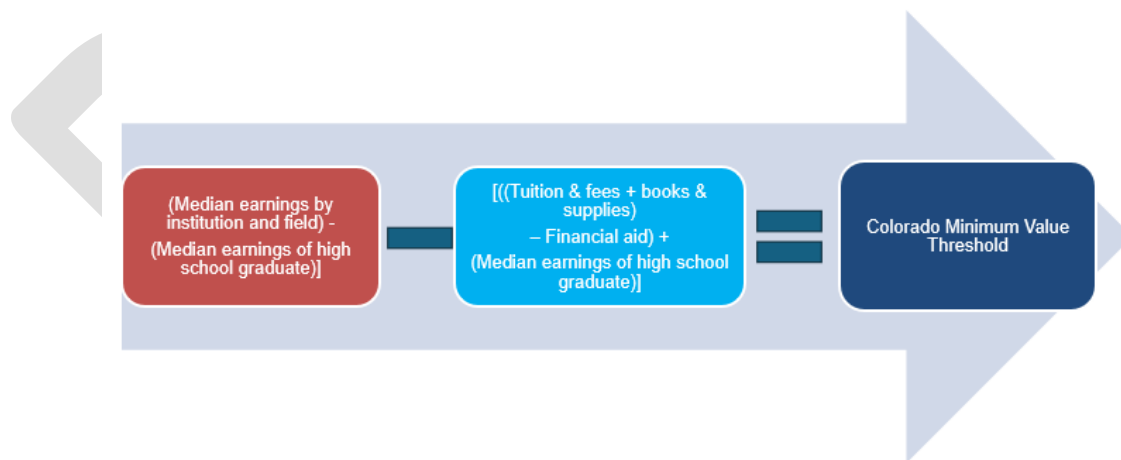
Realized earnings

We can observe former students' actual earnings through unemployment insurance records, which record the wages and salaries that an individual earned each quarter, if they lived in Colorado. If an individual was not employed or did not live in Colorado, their earnings are recorded as zero or missing. We can calculate annual earnings by summing the quarterly wages for each calendar year.

We should calculate annual earnings for individual-year observations with at least three quarters of positive earnings. Individual-year observations with fewer than three quarters of positive earnings should be treated as missing. This is because that individual could have been working out of state (meaning their earnings would not be recorded in the Colorado unemployment insurance records), so including this individual in the calculation would bias estimated earnings for their cohort downwards.

It is also possible to calculate realized earnings for students who work while they are enrolled in school, which may be especially applicable for students who work part-time to defray the costs of college and recoup some lost labor market earnings. Where these in-school earnings are available, they should be incorporated into the MVT calculation.

Where in-school earnings are missing, but enrollment data show that a student was enrolled at a Colorado institution of higher education, that student's earnings should be recorded as zero rather than missing. Or, if the student has fewer than three quarters of positive earnings, these partially missing earnings should still be recorded and included in the calculation. This is because the reason for the students' missing or partially missing earnings is almost certainly their enrollment in school, so we can reasonably conclude that their actual earnings during the missing quarters were zero.

Counterfactual earnings
Ideally, counterfactual earnings would reflect what a college student would have earned in the parallel universe where she did not attend college. However, we can settle for constructing a counterfactual earnings estimate based on the earnings of people who are similar to each college student at the point of college entry but have only a high school diploma.

We can use the American Community Survey (ACS) to calculate counterfactual earnings. The variable INCWAGE records each ACS respondent's wage and salary income over the past 12 months and is thus most comparable to the unemployment insurance wage and salary income recorded in Colorado's unemployment insurance records.

To construct an ACS sample that we can use for counterfactual estimates, we should include only ACS respondents who live in Colorado, have a high school diploma but no college credential (EDUCD between 63 and 65), worked for at least 27 weeks of the previous 52 (WKSWORK2 between 3 and 6), and had nonzero wage and salary income (as recorded by INCWAGE).

The SURDS enrollment file records, for each student: gender, race and ethnicity, date of birth and year of separation from the institution (from which we can calculate each student's age upon separation), and for some students, the Colorado county in which the student lived at the time of entry. ACS records gender, race and ethnicity, age, and county and metro area of residence.

We can therefore use ACS data to calculate counterfactual earnings for a demographically and geographically similar group of individuals to each student, using the student characteristics information recorded in the SURDS file. We divide the ACS sample into groups along the following dimensions: gender (male or female), race and ethnicity (underrepresented minority or non-underrepresented minority),

3

geographic location (Front Range urban corridor or rest of state), and age in 2007 or 2008 (16 to 21, 22 to 24, 25 to 29, 30 to 34, or 40 to 49).

We match each student in the SURDS file to the appropriate counterfactual group of individuals in ACS. For instance, a white female student who lived in Denver County upon enrollment and began college at age 18 in 2007 would be matched to the ACS subsample of non-underrepresented minority women who live in the Front Range urban corridor and were between 16 and 21 in 2007.

We can also use broader subsamples for student observations with missing data. If this example student lacked race or ethnicity information, for instance, we could instead use an ACS subsample of all women who live in the Front Range urban corridor and were between 16 and 21 in 2007.

We can then use this method to assign counterfactual earnings estimates to every student-year observation for which we can calculate realized earnings in the unemployment records. This example student, for instance, would have different counterfactual earnings in 2007, in 2008, in 2009, and all the way through until 2021, reflecting the fact that her counterfactual earnings profile rises over time.

College costs

CDHE records published tuition figures for each Colorado institution and program (along with the estimated cost of books and supplies), while the SURDS financial aid file records the scholarship and grant aid given to each student. We can subtract each student's financial aid package each year from the cost of tuition, fees, books, and supplies associated with their institution and program (adjusted by enrollment intensity). This allows us to construct a measure of net price for each student in each year that they are enrolled in school. We add up the present discounted value of the net price paid in each year to ascertain the total cost of college for each student. To aggregate costs, we take the central tendency of this total cost rather than the central tendency for each year.

Calculating minimum value threshold

All cashflows (realized earnings, counterfactual earnings, and tuition payments) are being used as nominal values. This table shows an example student who enrolls in college in 2008, stays for four years, then works in-state for seven of the next ten years:

Years since first enrollment (used to calculate discounting)	Calendar year	Student activity	Realized earnings	Counterfactual earnings	College costs
0	2007	In school (1/2 year)	None; included as zero	Half-year	Net cost (1/2 year)
1	2008	In school (full year)	None; included as zero	Full year	Net cost (full year)
2	2009	In school (full year)	None; included as zero	Full year	Net cost (full year)
3	2010	In school (full year)	None; included as zero	Full year	Net cost (full year)
4	2011	In school (1/2 year); working in-state (1/2 year)	Half-year; included	Full year	Net cost (1/2 year)
5	2012	Working in-state (full year)	Full year; included	Full year	None
6	2013	Working in-state (full year)	Full year; included	Full year	None
7	2014	Working in-state (full year)	Full year; included	Full year	None
8	2015	Working out-of-state (full year)	None recorded; dropped	Not calculated	None
9	2016	Working out-of-state (full year)	None recorded; dropped	Not calculated	None
10	2017	Working in-state (full year)	Full year; included	Full year	None
11	2018	Working in-state (full year)	Full year; included	Full year	None
12	2019	Employed less than 1/2 year	Only two quarters recorded; dropped	Not calculated	None
13	2020	Working in-state (full year)	Full year; included	Full year	None
14	2021	Working in-state (full year)	Full year; included	Full year	None

Because any individual student is likely to have several instances of missing data in the unemployment records, it is not advisable to calculate MVT for an individual student—since we cannot be sure whether that student is employed out of state, earns money through non-unemployment insurance sources, or is not employed at all. We should therefore drop missing instances of realized earnings rather than treating them as zero (unless the student is enrolled in college that year).

Instead, we should aggregate students into cohorts based on their year of entry into higher education. Within that cohort, we can then calculate the central tendencies (averages or medians) of realized earnings, counterfactual earnings, and college costs for each calendar year, dropping missing values. While individual student-year observations may be missing, each cohort is likely to have some observations in each calendar year. However, since the data is linked to individual students, we can assign students to multiple cohorts depending on the questions we want answered.

Proposed timeline

Date	Task
6/28	DAG meeting – in-depth conversation on model methodology and datasets.
Week of 7/1	Sharing updated technical document with added FAQs with stakeholder groups.
Week of 7/8	Sharing some of the datasets that will likely be used as part of the model for review and additional feedback.
Week of 7/15	Continued conversations with CFOs (7/16) and DAG (7/18) on technical documentation, FAQs, and datasets.
7/24 – 7/26	CCHE retreat – presentation on HB22-1349 implementation, minimum value threshold model, and discussion/feedback on next steps.
August	Sharing some preliminary model data with IHEs/governing boards. Potential additional meetings/presentations with CFOs and DAG. Continued discussions among CCHE with stakeholders on next steps.
September	Potential additional meetings/presentations with CFOs and DAG. Continued discussions among CCHE with stakeholders on next steps.
Mid-late September	Sharing of final model data with IHEs/CCHE for viewing and review. Holding collaborative conversations between CDHE, CCHE, and IHEs on findings and potential actions to increase student success.

Frequently Asked Questions (FAQs)

The following questions have been submitted to CDHE from various stakeholder groups (including but not limited to the Data Advisory Group, the Chief Financial Officers, and the Technical Working Group). CDHE staff has provided answers to those questions below. CDHE will continue to update these questions and answers as more feedback is provided.

Please submit additional questions via this Form:
https://forms.office.com/r/VNFtQmNKjc
CDHE will endeavor to add those questions and answers to this running list as this document continues to be updated.

General questions

Question	Answer
Enrolled students or degree recipients? – CDHE uses the term "student" but does not specify clearly… should be "degree recipients" or "award recipients"	Starting cohort of enrolled students. Model will show outcomes for completers and non-completers.
All students or just Colorado residents: Unemployment insurance wage and salary data are limited to Colorado so this analysis should focus on Colorado residents at time of entry. What's the impact of a high percentage of non-resident students have on these calculations?	Colorado in-state tuition and post school in-state wages
"Lifetime earnings" or time period for evaluating earnings outcomes	15 years of unemployment insurance earnings, additional years will be projected using 2% adjustments for both earnings and counterfactual data.
Why is 15 years being proposed for use instead of lifetime earnings?	Data availability limits CDHE's ability to calculate the full lifetime earnings (40+ years) for those in postsecondary education. 15 years of earnings data are available for current use, and CDHE can expand the years used as more data become available. Additionally, CDHE may work in future iterations of the model to extrapolate earnings growth after 15 years in an attempt to better reflect lifetime earnings.
What level of degree program detail will be used (2/4/6-digit CIP Code)?	Depending on available cohort size, 6-digit CIP if the cohort is big enough. If cell sizes are too small, CIP codes may be rolled-up to the 4-digit or 2-digit CIP level.

What geographic assumptions are being made (U.S., CO, MSA, county)?	Front Range Urban Corridor: Adams, Arapahoe, Boulder, Broomfield, Denver, Douglas, El Paso, Jefferson, Larimer, Pueblo, Teller, Weld Rest of Colorado: Alamosa, Archuleta, Baca, Bent, Chaffee, Cheyenne, Clear Creek, Conejos, Costilla, Crowley, Custer, Delta, Dolores, Eagle, Elbert, Fremont, Garfield, Gilpin, Grand, Gunnison, Hinsdale, Huerfano, Jackson, Kiowa, Kit Carson, La Plata, Lake, Las Animas, Lincoln, Logan, Mesa, Mineral, Moffat, Montezuma, Montrose, Morgan, Otero, Ouray, Park, Phillips, Pitkin, Prowers, Rio Blanco, Rio Grande, Routt, Saguache, San Juan, San Miguel, Sedgwick, Summit, Washington, Yuma
How does the region of high school affect earnings and counterfactual?	Counterfactual is assigned to the individual based on race/gender/geographic area, and the counterfactual is compiled from the population in the cohort.
What demographics are planned to evaluate? Are demographic details available for all elements of formula (graduate earnings, high school earnings, opportunity costs?	Demographics are based on the enrolled student body. Use of demographic detail is limited is when the cohort and/or cell sizes are 10 or less.
How will multiple awards be handled? Ex. If a student receives a certificate + degree, are earnings assigned to first award? Second? Both? How are stacked credentials (all levels) recognized?	Highest degree, limited to AS and BS will be used. Certificate data may be used but may be limited
Are graduate degrees included in the evaluation?	Not at this time. CDHE has access to data for this analysis and may be incorporated in future iterations of this work.
Will earnings for college graduates and high school graduates be sourced from the same data? Or mixing sources?	No, Colorado unemployment insurance data will be used for postsecondary graduates and American Community Survey for high school graduates.

"Median earnings" of college degree recipients

Question	Answer
Data source? CDLE? PSEO? National sources (BLS, USCEN)? Other?	Colorado unemployment insurance earnings

Age of earnings? All ages, 25-34? from graduation (year 1,5,10)?	Combined earnings for 15 years. We expect to be able to breakdown the data by the enrollment starting age.
Is there an inflation assumption (to grow earnings)?	No, nominal wages and tuition data being used.
Omissions from unemployment insurance wages and salaries: What populations are excluded from unemployment insurance wage records? How are work-study students, students on fellowship, and other forms of support recorded in these wage records? Also, college students could have summer employment opportunities outside of Colorado, so these wages are excluded from the analysis.	For enrolled students, any unemployment insurance data available will be used. Enrolled students will have all available data used. Non-completer students will require 3 of 4 quarters being present to be used.
Discount rate for present value calculation: Should use actual inflation rate to calculate real earnings per calendar year.	No, nominal wages and tuition data are being used.
Missing data for an individual student (middle of page 6): It's extremely important to get clarification on the inclusion of observations with partial earning data. Students are a highly mobile population so we should exclude observations when most of the years have missing data for an individual.	For non-enrolled, we are limiting data to 3 of 4 quarters of data to be used

"Median earnings" of high school graduates (aka Counterfactual)

Question	Answer
Data source? CDLE? Census? ACS? PSEO?	American Community Survey
ACS earnings are limited to half time or more (27 weeks plus) but is this comparable to unemployment insurance wages records?	Yes, we are using the "wksworked" field within American Community Survey data of 27 weeks or greater
We need an ACS translation table that defines the variables (e.g., EDUCD, WKSWORK2, etc.) and their associated variable values (e.g., EDUCD between 63 and 65). The following is based on information in Table S1501. https://data.census.gov/table?q=educational%20attainment Education (EDUCD) between 63 and 65 isn't explicitly stated in the table so I am recommending that the earnings are limited to "High school graduate (including equivalency)" We don't want to include "Some college, no degree" and "Associate's degree".	Correct. CDHE will continue to build out technical documentation on how datasets are pulled.
Age of earnings? All ages, 25-34, other?	All, should be broken down by cohort size.

Is there an inflation assumption (to grow earnings)?	No, nominal wages and tuition data being used.
Is income based on Age? County? Occupation?	Earnings data is based on age and county, but occupation data is limited and may not be used in the counterfactual.
HS Concurrent or not… how would this affect assumed earnings for non-college?	Not evaluated
Academic preparation? How would this affect control group earnings?	Academic preparation data are unavailable to us at this time

College Costs

Question	Answer
Data source?	Tuition and fees: Annual Tuition and Fees report and SURDS data. Books and supplies: IPEDS
What specific costs are included? Cost of Attendance less living expenses?	Tuition and fees and books and supplies
Will CDHE provide institutions a data set, on a student level, to validate tuition and costs allocated to a student in an academic year?	Yes. Data was provided for verification to the CFOs and feedback was incorporated.
What costs are allocated to students who are enrolled part-time?	Costs are calculated on a per credit basis
Will financial aid packages (grants & scholarships) be included in calc?	Yes
Which SURDS variables are included in the financial aid package calculation?	SURDS Financial Aid fields: Federal PELL, Federal SEOG, Other Federal Grants, CLEAP, CSG - CO Student Grant, CO Undergraduate Merit, CO Graduate Need Based, CO Graduate Merit, Filler - CO Categorical Grant, Filler - Inst Award Outside Funds, Inst Need Based Awards, Inst Merit Based Funds, Other Scholarship, Filler - Governors Opportunity Scholarship, SLEAP, GPA of Merit Recipient, Filler - Veterans Benefits, GEAR UP Scholarship, Filler - Academic Competitiveness, Filler - National SMART, Filler - CO PreCollegiate, CO Teach Scholarship, Federal Teach Scholarship, CO CTE Grant

Costs vary by degree type (cert, associates, bachelors), major, institution; Are certificate program costs available anywhere? There is a lot of variance in certificate costs by field of study and institution.	We are using base resident cost. Tuition differential rates are not always clear on the program and are not assigned to CIP codes.
Are there year assumptions by award level, such as 1 year for certificate, 2 for associates, 4 for bachelors? Are credits earned during high school built into the formula, for those that bring credits and shorten time-to-degree?	We are calculating costs by tuition paid for after initial enrollment. Transferable credits earned prior to enrollment should reduce cost by reducing the required credits.
From CDHE document: Central tendency of this total cost rather than the central tendency for each year: Is this value the median for degree recipients by degree in a calendar year? Is this value calculated based a student's institution, residency, degree, and major?	As we follow each student in the education path, we will include the cost of all credits earned for that student. When the data are summarized, we will take the median cost for the cohort. Cost will be base tuition rates for the in-state at the institution they take credits from.

"Opportunity Cost," earnings of high school graduates "lost" while enrolled in college

Question	Answer
Data source?	American Community Survey
Age assumption of earnings during college? 18-24? Actuals from CDE HS Graduates + CDLE? BLS?	Some disaggregation by age may be available, however cohort size constrains may limit their use
Is the assumption that students don't work during college?	Wages earned during college years will be included.
Similar question with college costs— what are the assumptions for years of college (for each degree type)?	Actual data on years and credits enrolled will be used, with no assumptions made.

Final Product Deliverable

Question	Answer
How long will institutions be able to review the tool (and data) prior to the tool's publication?	Data will not be public; no public tool at this time. Data will be provided via Excel to individual IHEs (potentially to governing boards) in August. Feedback will be solicited from IHEs as plans are made for collaborative conversations among CDHE, CCHE, and IHEs.
Has there been discussion about how to present the data findings?	Yes, we have been asked to present each institution/program/demographic as green or yellow with no

	numbers. We are reviewing option to provide additional data to each institution directly
Will it be a simple Yes/No for a program, dollar amount, or number of years?	The model indicates dollars
Will each element in the formula be available (both earnings, costs, inflation, all facets of calc)	Data elements will be provided to IHEs as part of the output calculations. Some PII/data privacy concerns will need to be balanced when sharing small cell sizes back with IHEs (since there are connections to wage data).
Will institutions have access to all statewide data and calculations?	IHEs will be provided with their own data, and a summary statewide will be provided to everyone
What is the chosen platform (ex Tableau)?	Direct communication with IHEs and data expected to be provided in Excel.
Will CDHE Staff or 1349 consultants develop and maintain the data model & visualization?	CDHE staff will likely maintain the model and make updates as needed. However, repeated evaluation of the same cohort and adding one year of earnings will have limited value.

Institutional control group

Question	Answer
Will each institution have specific control group?	No
How balancing multiple input demographics for control groups?	The enrollment cohort will have demographic information in the enrollment file that we will be using.
How to determine income of HS pop? FRL status? Household AGI?	American Community Survey