American Journal of Business and Operations Research

♦ ISSN: 2692-2967 Vol: 14, No. 01, 2026 ♦ PP. 04–09

Article DOI: https://doi.org/10.54216/AJBOR.140102



FULL LENGTH ARTICLE

Business Intelligence for the Global University Market: Bounding Recruitment Concentration and Regional Spending Exposure

Nilufar Ismailova[1,*]

[1] Department of World Economy and International Economic Relations, Tashkent State University of Economics, Tashkent, 100003, Uzbekistan

Email: nilufar.ismailova@tsue.uz

Received: August 02, 2025 Revised: October 12, 2025 Accepted: December 10, 2025 ⋆ Corresponding author

ABSTRACT

International recruitment exposes universities and host regions to concentrated cross-border demand, yet published dashboards rarely disclose the joint distribution of source countries and institutions. This study turns the November 2024 *Open Doors* state fact sheets into a reproducible market-risk screen for 50 U.S. states and the District of Columbia. These markets contain 1,126,079 international students and $43.817 billion in estimated regional expenditure in 2023/24. The median state's largest source market accounts for 28.5% of students; its largest receiving institution accounts for 32.2%. Sharp bounds calculated from the separately published margins show that, in 48 of 51 markets, a positive share must both come from one of the five leading origins and attend one of the leading institutions. The median guaranteed overlap is 32.0% of a state's international students; Arizona's is 68.4%. An illustrative 10% decline in each state's largest origin implies $1.511 billion in expenditure exposure across the 51 markets under a constant-spending assumption, led by California ($233.1 million) and New York ($229.9 million). The bounds identify segments requiring institutional cross-tabs; the scenario is neither a forecast nor a measure of university revenue. The analysis supplies an auditable way to prioritize recruitment diversification and market-level due diligence with public business data.

Keywords: Business intelligence ▪ International services ▪ Market concentration ▪ Partial identification ▪ Scenario analysis ▪ Higher education institutions

1. INTRODUCTION

Universities operate in a cross-border market for students, while local firms and households participate in the associated flows of spending. Leaders deciding where to recruit, build partnerships, or reserve contingency funds need to see both the scale of a market and its concentration. A high-volume state can have modest proportional dependence on one source country, while a smaller market can rely heavily on one intake channel. Treating each as the same type of risk obscures the decision.

The available public information presents a sharper problem. The 2024 *Open Doors* state fact sheets publish international student totals, leading origin shares, leading institutions, and estimated regional expenditure on separate panels [1]. They do not show the cross-tabulation of origins by receiving institutions. Managers therefore cannot calculate directly how many students at a given institution come from a given market. We ask what a decision analyst can establish *without* inventing the missing joint data, and how a transparent sourcing scenario changes the ranking of markets for investigation.

© 2026 The Author(s). Published by ASPG. This article is licensed under CC BY 4.0.

Research on international mobility cautions against equating student flows with a uniform revenue strategy: the association between funding patterns and international enrollment differs across systems [2]. Quantitative benchmarking of mobility has been developed for countries and networks [3]; business analytics research emphasizes the value of data quality and decision routines, rather than prediction for its own sake [4]. We contribute a compact, auditable portfolio screen for a specific business decision. Our outcomes are exposure indicators for cross-border services and regional expenditure. They are not measures of teaching quality or student attainment.

2. DATA AND BUSINESS MEASURES

2.1 Published market data

We use the 2024 state fact sheets released by the Institute of International Education and the U.S. Department of State in November 2024 [1]. The student counts and origin and institution rankings concern academic year 2023/24. The fact sheets attribute estimated international student expenditure to NAFSA's economic analysis, which uses *Open Doors* enrollment data; the broader estimate covers spending beyond a university's accounts [5]. We take all 50 states and Washington, D.C. as the market frame, omitting Puerto Rico and the U.S. Virgin Islands. Their 611 students explain the difference between our 1,126,079 and the national 1,126,690 reported in *Open Doors*. No post-June 2025 data or literature enter this study.

For state i, let N_i be its reported international-student count, E_i its estimated expenditure in U.S. dollars, o_i the published share from its largest origin, and c_i the largest listed institution's count divided by N_i. The fact sheets list up to five origins and five institutions. Let O_i be the sum of listed origin shares and C_i the sum of listed institution counts divided by N_i. The denominators in o_i, O_i, c_i, C_i are the *same state's student total*; the expenditure measure has a different unit. Reported origin shares have one decimal place, so overlap percentages inherit their rounding. We preserve underlying institution counts and expenditures without imputing unlisted countries or institutions.

The market-size screen compares E_i with o_i, and the institutional screen compares O_i with C_i. These answer different commercial questions: which region has the greatest spending scale exposed to a single foreign intake, and where could a disruption reach concentrated institutional operations? The spending estimate is regional activity including purchases outside institutions. It must not be labeled tuition, gross institutional receipts, profit, or export earnings directly observed in these tables.

2.2 Identification from marginal totals

Define A_i as the set of students from a state's listed leading origins and B_i as the set enrolled at its listed leading institutions. Without a cross-tab, the share q_i in both groups is unidentified but has sharp logical bounds:

$$\max\{0, O_i + C_i - 1\} \le q_i \le \min\{O_i, C_i\}. \tag{1}$$

The lower bound follows because two sets whose shares sum to more than one must overlap by at least their excess over one. Either endpoint can be attained by an allocation consistent with the two margins; hence these are identification bounds rather than confidence intervals. We repeat the calculation with o_i, c_i for the single largest origin and institution. The leading-institution list has fewer than five entries in some small states; B_i always means the *listed* institutions. Where the sum $O_i + C_i$ does not exceed one, zero guaranteed overlap does not imply zero actual overlap.

A useful operational reading of Eq. (1) is to imagine laying the two ranked lists over the same population. The sum of their shares may exceed 100%, in which case the excess is the number of students who must belong to both lists, expressed as a fraction of the state total. Nothing in the fact sheets says how that excess is allocated among individual origin-country and campus cells. In a portfolio with five origins and five institutions, a positive group-level lower bound could be spread over many of the 25 named cells. The formula therefore supports a guaranteed *group* exposure, not a claim that one institution is dependent on one country. The upper bound represents the other extreme: the smaller group could fall entirely inside the larger. This range makes the information gap visible without assuming independence, random matching, or a uniform distribution of students across campuses.

2.3 A scale-sensitive stress scenario

For comparison with relative concentration, suppose each state's leading origin supplies 10% fewer students, all other counts are unchanged, and expenditure per enrolled student remains constant. The arithmetic exposure proxy is

$$S_i(0.10) = 0.10\, o_i E_i. \tag{2}$$

We sum S_i across the 51 markets. The exercise independently shocks the leading origin of *each* state, which may be a different country. It is a fixed-share scenario, not a probability-weighted loss, causal estimate, or forecast. Actual spending responses could differ with tuition, program mix, duration of stay, household expenditure, replacement enrollment, and local multipliers. The equation deliberately uses no fitted parameters.

An alternative scenario asks what a shock to one *named country* would look like across the U.S. portfolio. For country k listed among a state's five origins, define its published share as p_{ik} and calculate $0.10 \sum_{i:k \text{ listed}} p_{ik} E_i$. States in which k is not listed contribute zero to this reported-share measure. This gives a lower bound on the scenario under the constant per-student-spending convention, because an unlisted country's true share can be positive. It should not be combined with the largest-origin scenario: one fixes the origin name across states, whereas the other changes the origin name to the leader in each state. This distinction matters for a multinational recruitment strategy whose partnerships and brand reputation depend on a particular foreign market.

Each of the 51 states and D.C. receives the same weight in reported medians, while sums of student counts and expenditures retain their own units. The exercise is descriptive over the observed portfolio of markets; we do not report a sampling confidence interval for deterministic logical bounds. A national average of state medians would have no coherent interpretation as the student-weighted joint overlap. The supplementary CSV preserves inputs for weights appropriate to

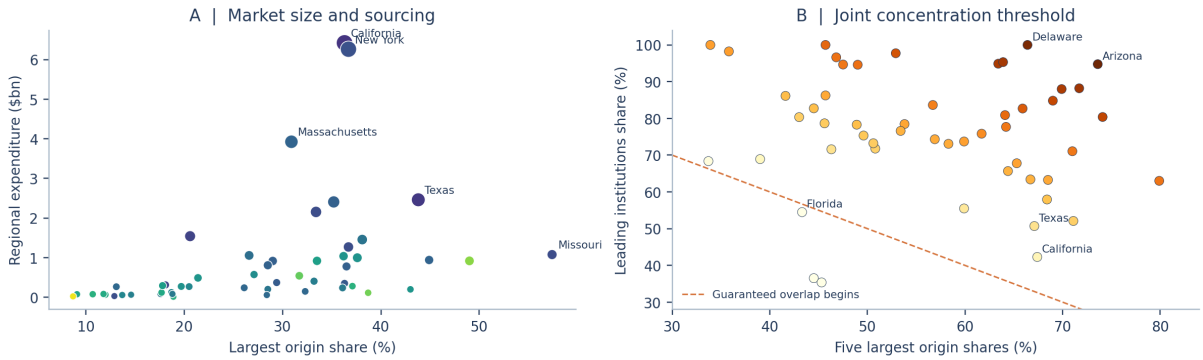


Figure 1. Two views of international recruitment markets, 2023/24. Marker size in panel A reflects total international students; marker color reflects the leading institution's enrollment share. Panel B plots separately published five-origin and leading-institution margins; above the dashed line their overlap has a strictly positive lower bound. States and D.C. are the 51 observations.

Table 1. Market portfolio summary for the 51 U.S. state and D.C. markets

Indicator	Value
International students	1,126,079
Estimated regional expenditure	$43.817bn
Largest origin, median state share	28.5%
Largest institution, median state share	32.2%
Listed origins, median state share	56.9%
Listed institutions, median state share	77.7%
Guaranteed joint listed share, median	32.0%
States with positive joint lower bound	48 of 51

Note. Medians are unweighted across 51 markets; they are not national shares. Expenditure is the sum of NAFSA-derived state estimates.

a particular organization's exposure.

The published origin shares are rounded to one decimal place, so a sum of five shares may deviate from the unrounded total by no more than 0.25 percentage points in either direction under ordinary rounding. Campus counts are integer totals. Consequently a positive five-origin overlap bound more than 0.25 percentage points above zero cannot disappear solely because of origin-share rounding. The smallest positive bound in these data is West Virginia's 2.1%, so the count of 48 positive bounds is stable to that rounding convention. The reported lower-bound percentages retain their calculated precision in the accompanying CSV to facilitate reproducibility, while the paper rounds them to one decimal place for interpretation. The scenario's leading-origin share differs by at most 0.05 percentage points under the same convention, much less than the differences between the major dollar rankings.

3. RESULTS

3.1 Size and concentration are distinct

Across the 51 markets, the largest origin supplies between 8.7% and 57.4% of international students (median 28.5%); the largest institution hosts between 12.4% and 77.8% (median 32.2%). Missouri has the highest leading-origin share, 57.4% from India, but its leading institution hosts only 18.6% of the state's international students. California's largest origin is China at 36.3%, while its leading institution hosts 12.4%. Arizona combines a 49.0% India share with a 66.1% leading-institution share. Figure 1 places these three very different portfolio shapes on the same scale. A large origin share therefore need not imply equivalent concentration at a single institution.

The estimated $43.817 billion in expenditure is highly uneven across the geographic portfolio. California contributes $6.421 billion and New York $6.263 billion; together with Massachusetts, Texas, and Illinois, these five markets account for 49.0% of the 51-market total. That concentration concerns *where spending occurs*, not the distribution of revenue across individual universities. The median state's implied estimated expenditure per student is $34,392; because the numerator is a model-based regional estimate, this ratio should not be treated as an individual student's invoice or average tuition.

Missouri has 32,647 international students and $1.076 billion of estimated regional expenditure. New York has 135,813 students and $6.263 billion of expenditure, with a lower largest-origin share. Texas has 89,546 international students but a 13.3% leading-institution share. These contrasts show why a recruitment department's origin risk and a state economic-development office's spending exposure should sit on separate dashboards even when both begin with the same enrollment source.

3.2 What the published margins guarantee

The median state has 56.9% of international students in its five listed origins and 77.7% at its listed leading institutions. Applying Eq. (1) state by state yields a median lower bound of 32.0% for their intersection. The lower bound exceeds zero in 48 markets. The three with a zero guarantee are Florida, Kentucky, and Montana; the missing origin-by-campus cross-tabs could nevertheless contain considerable intersection there.

Arizona illustrates the operational interpretation. Its five largest origin shares sum to 73.6%; its five listed institutions account for 94.8% of its 27,883 international students. At least 68.4% of the state's students must belong to *both* groups, while as many as 73.6% could. For California the corresponding margins are 67.4% and 42.3%, giving an interval of 9.7–42.3%. The considerably wider California interval signals that the public tables do not resolve which particular campus and origin combinations dominate. Figure 2B also shows why a state should not be assigned the midpoint of its feasible interval as if it were an observed segment share.

Delaware's five listed institutions together account for all 4,023 reported international students, so its five-origin group

is wholly contained in that institutional group: the lower and upper overlap bounds coincide at 66.4%. In Florida the five-origin share and five-institution share sum to at most 100%, leaving a zero lower bound even though the sets may have members in common. These results are mathematical properties of published margins, not evidence that Delaware has a safer revenue portfolio or Florida has no common recruitment channels. A manager seeking a single institution's exposure still needs the unreported cross-tab.

The largest-origin and largest-institution shares have a weak negative rank association across markets (Spearman's $\rho = -0.274$). The small number of markets and differences in institutional grouping limit interpretation, but one cannot substitute one concentration dimension for the other. A large state can have students from one country distributed across many campuses; a small state can have diverse source countries concentrated at one institution. Separate marginals contain useful but incomplete signals for these different arrangements.

The largest-single-origin/largest-single-institution screen is even more conservative. Its guaranteed overlap is positive only in Arizona (15.1% of all international students) and Delaware (1.5%); in the other 49 markets the marginals alone do not force a positive intersection. That result does not imply that the largest campus receives no students from the largest origin. It reveals the limits of marginal dashboards and directs institutions toward the origin-by-campus counts needed for a narrower assessment.

The market-screen table clarifies that the same origin concentration can coexist with quite different levels of guaranteed group overlap. Maryland and California have almost identical largest-origin shares (36.2% and 36.3%), yet their listed-group lower bounds are 59.9% and 9.7%. The principal difference in these two margins is the share housed at listed institutions, 88.2% in Maryland and 42.3% in California. Their estimated expenditures also differ greatly. A business unit choosing which state to investigate can thus make three separate choices: prioritize origin concentration, prioritize the proven intersection with major institutions, or prioritize dollars. Without specifying the decision, a single composite "internationalization risk" score would conceal the weights and likely change the ranking arbitrarily.

There is no contradiction between the national-scale student total and the unweighted median state. The first aggregates persons across geographic markets; the second describes the typical market as one equally weighted observation. A consortium with campuses concentrated in a few states would obtain different exposure totals by applying its own footprint. Nor does the guaranteed overlap of 32.0% in the median state mean that precisely 32.0% of U.S. students occupy a particular origin-institution combination. It is a median of conservative group bounds, each formed from that state's published margins.

3.3 Dollar exposure under a common shock

Equation (2) gives $1.511 billion in summed expenditure exposure, 3.45% of the baseline across the 51 markets. The leading cases are California ($233.1 million), New York ($229.9 million), Massachusetts ($121.3 million), Texas ($107.8 million), and Illinois ($84.7 million). Missouri, despite its highest leading-origin *percentage*, ranks sixth by dollars at $61.8 million. Conversely, New York's leading-origin share is 36.7%, yet its large $6.263 billion spending baseline places it second by scenario dollars. Figure 2A makes the ranking change visible.

Table 2. Contrasting market exposure signals

Market	Origin (%)	Overlap (%)	Stress ($m)
Arizona	49.0	68.4	45.0
California	36.3	9.7	233.1
Maryland	36.2	59.9	37.5
Missouri	57.4	42.9	61.8
New York	36.7	23.2	229.9
Texas	43.8	17.8	107.8

Note. Joint minimum is for listed leading origins intersecting listed leading institutions. Scenario dollars are hypothetical regional exposure; rows are illustrative selections.

The scenario is linear by construction. Halving the assumed proportional decline to 5% halves the exposure to roughly $0.756 billion; raising it to 15% yields about $2.267 billion, under the same assumptions. Such sensitivity demonstrates arithmetic, not validation against a later observed shock. A university planning a contingency budget should replace the state average with its own origin-by-program enrollment, net tuition, aid, retention, and variable-cost figures.

Holding the affected country fixed changes the interpretation. India appears among the five published origins in 50 of 51 markets. A 10% reduction in its listed shares corresponds to at least $1.251 billion of modeled regional expenditure under the constant-spending convention; China appears in 48 markets and gives at least $1.136 billion. These figures include states where the named origin is second through fifth, unlike the $1.511 billion leading-origin scenario. In the India screen New York contributes $149.7 million and California $128.4 million; in the China screen the same states contribute $229.9 million and $233.1 million. For China the top two markets account for 40.7% of the reported-share total. A geographically diversified U.S. portfolio can therefore share a common foreign-demand exposure. We neither add the two country totals as a single scenario nor impute country shares where they were not reported in the top five.

4. MANAGERIAL INTERPRETATION

The results support a two-stage workflow. A regional development office or university consortium can use dollar exposure to rank where a foreign-market slowdown may have the largest economic footprint, then use the joint lower bound to decide where limited origin-by-institution detail would be most valuable. The largest-origin percentage alone would prioritize Missouri before California; a scale-sensitive view reverses that comparison. A guaranteed large joint segment, as in Arizona, merits tracing specific intake pipelines and contingency capacity across the listed institutions. In California the broad bound chiefly calls for a cross-tab before assigning responsibility to any one campus.

A regional employer coalition considering housing, transport, or internship capacity may start with the expenditure ranking and then investigate which cities and programs actually attract the students. A university finance team needs a different

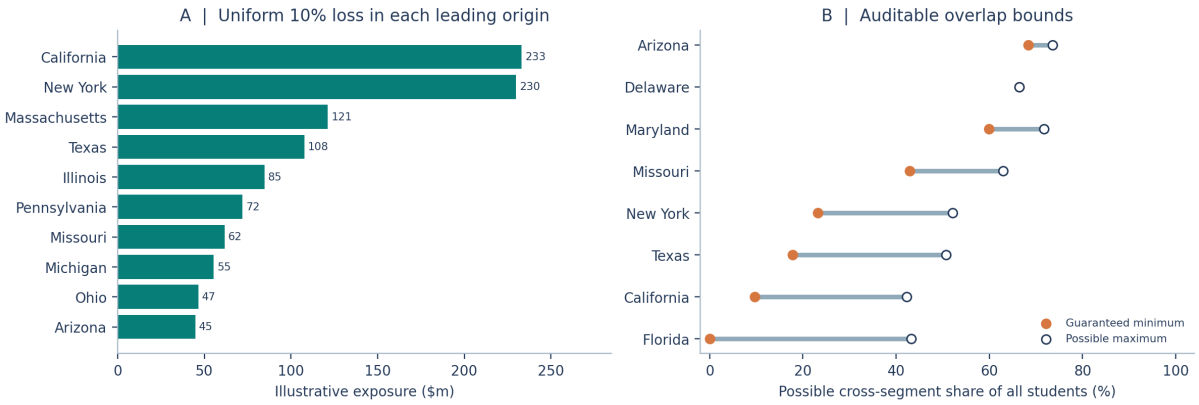


Figure 2. Two decision aids from the same published data. Panel A ranks the ten largest dollar exposures under a simultaneous hypothetical 10% fall in each state's own largest origin; values are regional expenditure proxies, not institutional losses. Panel B shows sharp logical intervals for the overlap of the listed leading origins and institutions in selected markets. Closed dots denote guaranteed minima and open dots possible maxima.

ledger: paid tuition net of waivers, scholarships, support costs, and the timing of receipts. Its recruitment team can reconcile lead-to-enrollment conversion by country with admissions data and verify whether a foreign-market slowdown can be offset by other qualified applicants. State agencies cannot observe these facts from the fact sheets alone. Each user should connect the public screen to its own records before attaching a cash-flow or staffing consequence.

Bounds also organize a request for more data. Where both marginals are high, ask the reporting institutions for a confidential origin-by-campus matrix. Where the lower bound is zero but the upper bound large, even a few verified cell counts can narrow uncertainty. When the endpoints coincide, as in the Delaware listed-groups example, collecting the same joint aggregate adds little to that question, though individual-campus decisions still need further detail. This is a value-of-information argument based on feasible intervals, not an assumed statistical distribution within them. It helps target analytics work without assigning students to unobserved cells.

This distinction matters to business intelligence. Even an accurate dashboard of separate marginals can support a misleading joint claim. The bound supplies a defensible minimum and maximum, while the scenario states a hypothetical intervention in explicit units. Together they make a decision more reviewable than a synthetic point estimate of campus-country exposure. As Szukits and Móricz [4] find in firms, data quality and analytical culture shape the usefulness of data for decisions; a more elaborate predictive model cannot recover a joint distribution that the public data never report.

The named-country scenario is useful when evaluating centralized recruitment operations. India is present in almost every state's leading-origin list, yet the largest Indian-dollar exposures are not simply the states with the highest India percentage: New York's scale puts it first in this screen. China's two largest listed-dollar exposures are California and New York, whereas a state office looking only at its own largest origin would not see the potential co-movement of those states. Even the named-country figures remain geographic spending proxies. To estimate a university's actual revenue-at-risk, the analyst must stratify by fee status and program, apply anticipated yield and continuation rates, and separate fixed costs from costs that change when a student does not enroll. Nothing in the public release can perform those steps.

The paper's U.S. application illustrates an information structure found in many international services markets: one table reports demand by customer origin, another reports demand by delivery site, and neither reports the crossing. Bounding the overlap is portable to those settings when the two margins genuinely refer to the same customer population and period. Here that requirement is met for the state-level student shares, but it would fail if we substituted NAFSA expenditure for a headcount margin or mixed different academic years. Maintaining common denominators is more consequential than choosing a complicated statistical model.

The study has several limits. State estimates are a single cross section, and we do not have the underlying institutional survey records, expenditure model variance, or cross-tabs. Rounding of source-market percentages slightly affects the calculated bounds. The expenditure series comes from NAFSA's economic model, so student counts and dollars do not constitute two independent confirmations of the same mechanism. The leading-origin stress assumes uniform per-student spending within each state; students from different countries and institutions can spend very different amounts. Nothing here estimates that recruitment effort causes a rise in enrollment, that a decline will occur, or that institutions themselves lose the reported dollars. International mobility also serves public and scholarly purposes beyond financial flows [2]; this paper deliberately confines its inference to commercial exposure in a U.S. market snapshot.

4.1 Decision use and next measurements

A practical dashboard should retain the raw denominators next to every percentage. For the origin panel it needs the country's name, its share of international students, the enrollment denominator, and the date of the source release. For the institution panel it needs the leading institution's count and the coverage of the listed group. For the economic panel it needs the NAFSA-derived expenditure estimate, its 2023/24 reference period, and the explicit scenario rate. Showing these fields together lets an executive see whether a high percentage reflects a large business footprint or simply a small denominator. Alaska, for example, has only 297 interna-

tional students; a large percentage movement there cannot be interpreted in dollars like the same movement in California.

A priority queue can then be constructed without claiming a statistically estimated risk of disruption. First sort by $S_i(0.10)$ if the mandate is to protect regional spending, or by o_i if the mandate is to diversify recruitment channels. Within that queue inspect the width $\min(O_i, C_i) - \max(0, O_i + C_i - 1)$ of the feasible origin-institution interval. A wide interval means the public release leaves much of the overlap unresolved and a requested cross-tab may change the assessment substantially. A narrow interval means the group total is already fairly well constrained, although the allocation among individual institutions can still be unknown. The ranking criterion and the interval width should remain separate dashboard columns. Combining them into a single index would silently impose a trade-off between dollars, concentration, and missing information that the decision maker has not specified.

Once institution-level data are available, the same framework can be checked rather than assumed. Compare the summed origin-by-campus cells with each published state total, reconcile source-country definitions and reporting periods, and measure the share of each institution's international enrollment from each origin. Link those cells to net tuition and variable support costs for an institution-specific margin scenario. If changes in applications are being forecast, use historical conversion and retention cohorts with a dated holdout period; a single 2023/24 cross section cannot calibrate either a trend or an early-warning threshold. The fact sheets are suitable for discovery and prioritization, and their greatest commercial value may lie in revealing which missing cells would alter an actual operating decision.

This decision sequence also clarifies responsibility for validation. The state or regional analyst owns the public spending estimate and can monitor local accommodation and consumer services. A university's finance office owns its fee and cost data; its international recruitment office owns country-specific applications and offers. Their data can be joined only with consistent institutional identifiers, geography, and enrollment dates. If an institution spans state lines or the list groups campuses differently from internal reports, those entities must be reconciled before a campus-level risk number is issued. Such checks do not change the state-level bounds, but they prevent the bound from being repackaged as a precision estimate for a unit it was never designed to measure.

The economic scenario should also be read at the level of an accounting identity. Applying a 10% change to an origin share while holding expenditure per student fixed scales a state total; it does not estimate demand elasticity or allow for substitution by applicants from another country. The country-specific totals based on five listed origins omit every unlisted state-country share. Their "at least" interpretation holds for nonnegative estimated spending under that same allocation assumption, not as a lower bound on an eventual real-world loss. A manager can use the scenarios to ask whether a possible range is material enough to warrant more data, while keeping observed 2023/24 expenditure separate from conditional future revenue. The distinction is essential when the organization carrying the recruitment risk is a university but the published dollar measure covers the wider host economy.

5. CONCLUSION

Public 2024 market tables support more than a ranking of international-student volumes. For 48 of 51 U.S. state and D.C. markets, the leading origin and institution groups must share a positive student segment, with a median guaranteed overlap of 32.0%. A common 10% leading-origin scenario produces $1.511 billion in illustrative regional expenditure exposure, dominated in dollars by California and New York. These figures provide a disciplined business screen for sourcing and contingency decisions while making explicit which campus-level and financial facts remain unobserved.

DATA AVAILABILITY

The original 2024 *Open Doors* state fact sheets are available at https://opendoorsdata.org/wp-content/uploads/2024/11/OD24_Fact-Sheets_States.pdf. The reproducibility package contains that source, the 51-market analysis CSV, extraction and plotting scripts, and both PNG figures.

REFERENCES

[1] Institute of International Education, "State facts and figures 2024," Institute of International Education, State fact sheets, 2024.

[2] T. Weber, C. van Mol, and M. H. J. Wolbers, "International students as sources of income? moving beyond the neoliberal framing of internationalization," *Journal of Studies in International Education*, vol. 28, no. 4, pp. 526–545, 2024.

[3] K. Breznik, M. Restaino, M. P. Vitale, and G. Ragozini, "Analyzing countries' performances within the international student mobility program over time," *Annals of Operations Research*, vol. 342, pp. 1925–1943, 2024.

[4] Á. Szukits and P. Móricz, "Towards data-driven decision making: The role of analytical culture and centralization efforts," *Review of Managerial Science*, vol. 18, pp. 2849–2887, 2024.

[5] NAFSA: Association of International Educators, "International students contribute record-breaking level of spending and 378,000 jobs to the U.S. economy," Nov. 2024.