REMI User Conference 2025, Ann Arbor, MI

Clustering Occupations by Knowledge

Methods and Insights

Presented by Indraneel Kumar, Ph.D., AICP

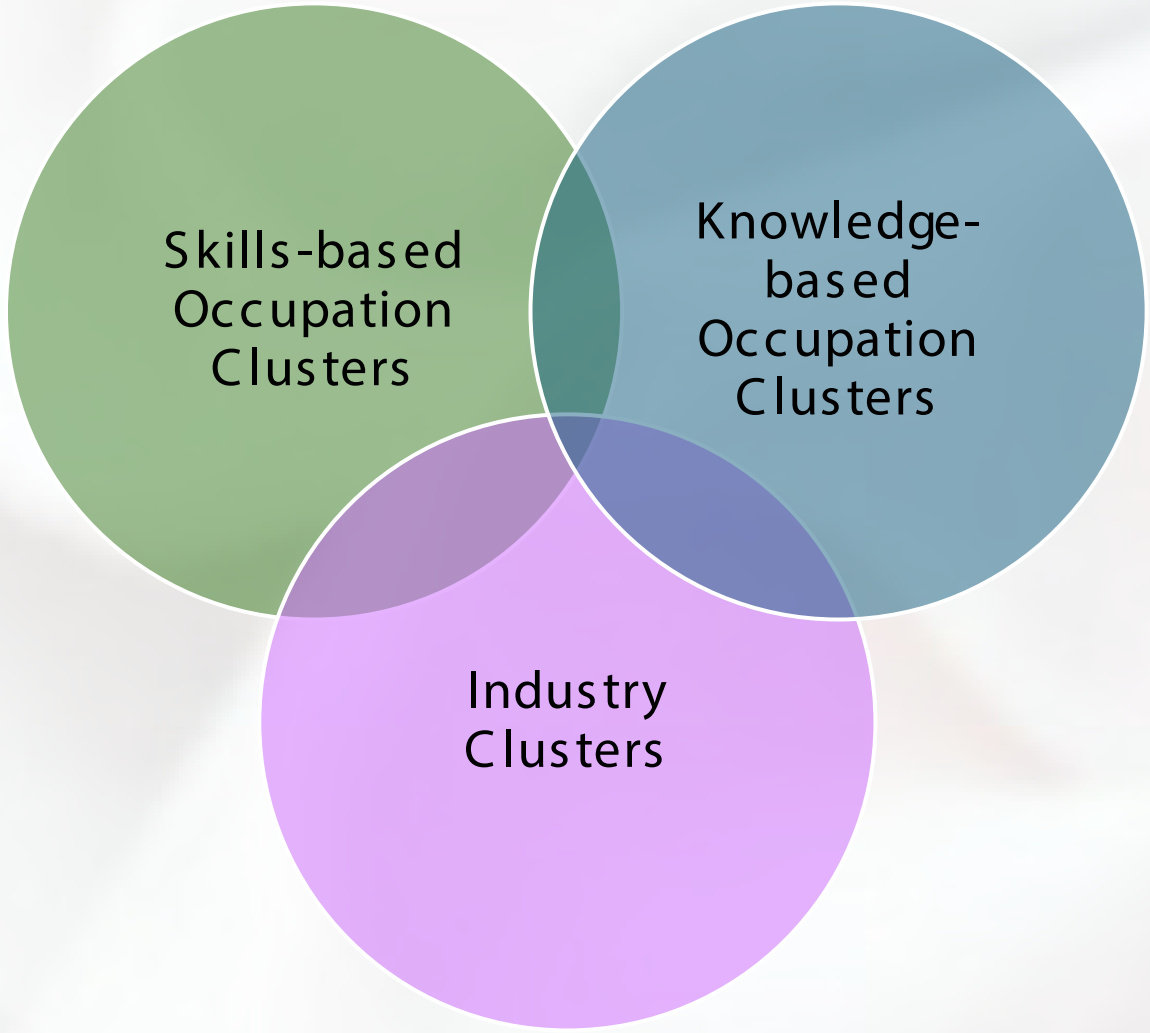
Disclaimer: This presentation is prepared by Purdue Center for Regional Development using Federal funds under award # ED23RNA0G0174 from the Economic Development Administration, U.S. Department of Commerce. The statements, findings, conclusions, and recommendations are those of the author(s) and do not necessarily reflect the views of Economic Development Administration or the U.S. Department of Commerce.

Center for Regional Development

October 2025

Economic Clusters Project Overview:

- 38 Skills-based occupation clusters, covering all 857 occupations in the U.S. labor market
- 21 Knowledge-based occupation clusters, covering 64% of total occupations requiring above high-school education.
- Approximately 25 primary industry clusters with 65+ subclusters.



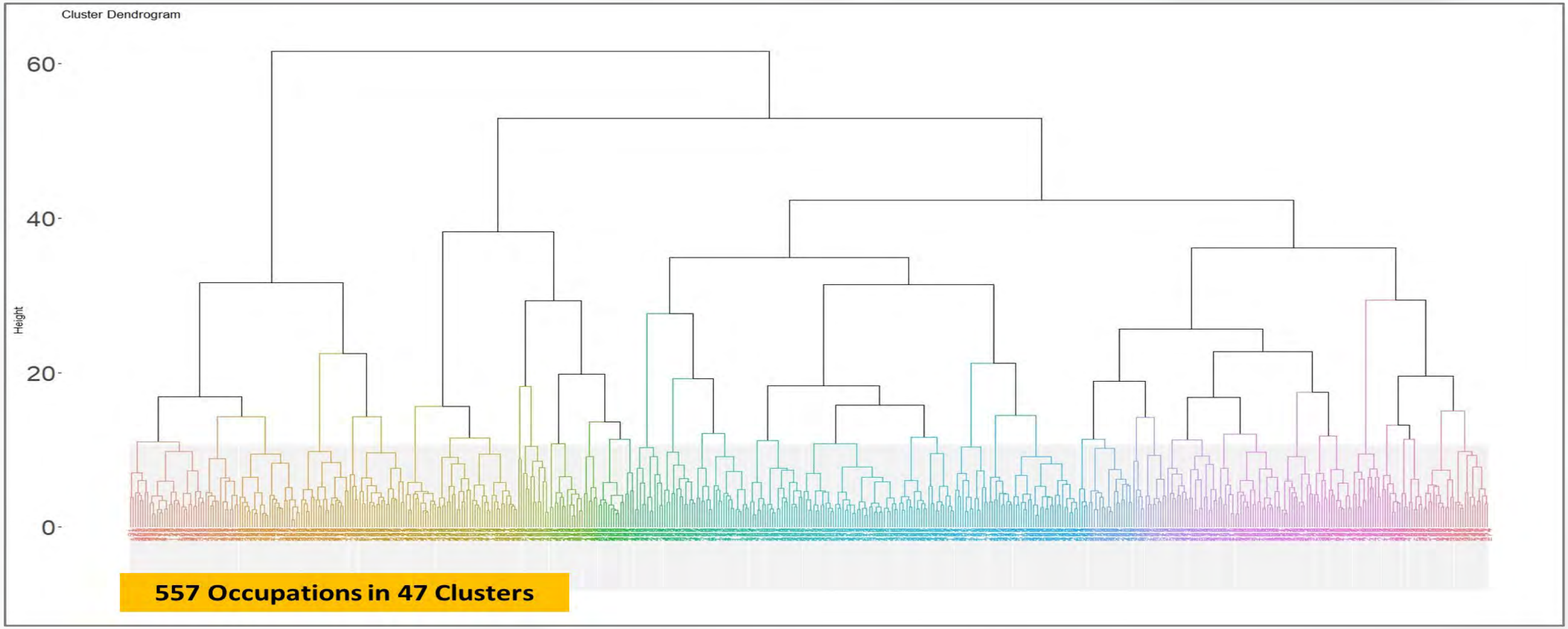
Knowledge-cluster Process Overview

- The project follows a mixed-method approach (qualitative and quantitative).
- Interviews of experts, literature review, data analysis, etc.
- Clustering method: (1) Partitioning (2) Hierarchical
- Partitioning methods can divide n objects into k groups, but user needs to provide k.
- Hierarchical methods explore different configurations of k, and either user can input k or choose k based on statistical specification tests.
- Partitioning: Partitioning Around Medoids (PAM), Clustering Large Applications (CLARA)
- Hierarchical: Agglomerative Nesting (AGNES), Divisive Analysis (DIANA)

$$\textit{Euclidean}(\textit{distance}) + \textit{Ward}(\textit{linkage}) + \textit{AGNES}(\textit{clustering}) + \textit{Silhouette Score (statistical specification)}$$

Knowledge-clusters Preliminary Results

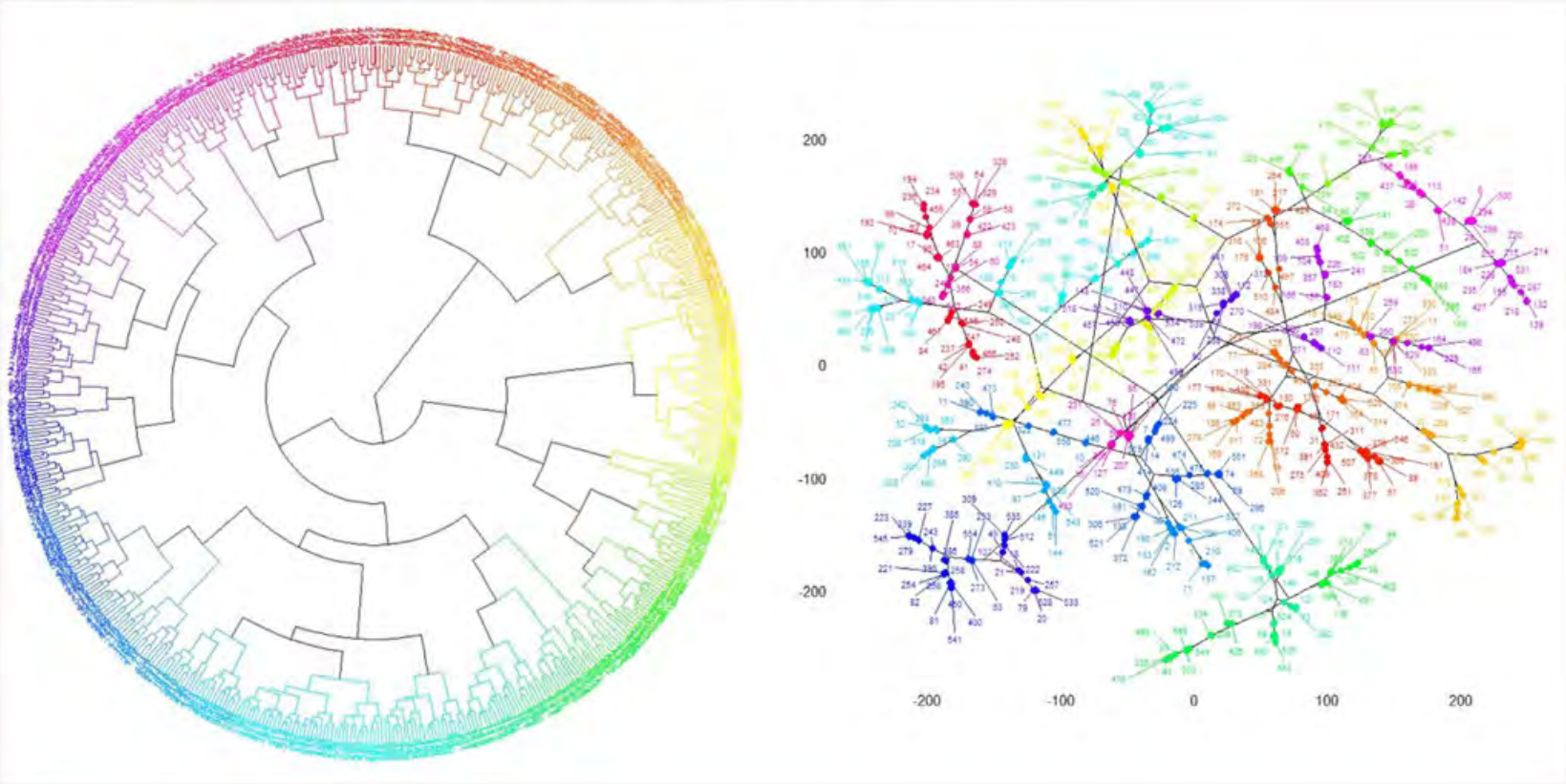
Result: Dendrogram of 47 Occupation Clusters (Knowledge-levels squared) Job Zones 3, 4 and 5



Knowledge-clusters Preliminary Results

- Specification tests were suggesting for 47 occupation clusters
- Different statistical tests included Elbow, Average Silhouette, and Gap Statistic methods

FIGURE 7. Circular and Tree Dendrograms of 47 Occupation Clusters for 557 Occupations



Note: Knowledge-levels squared; Job Zones 3, 4 and 5.

Source: Based on Gap Statistic Method. Processed by PCRD.

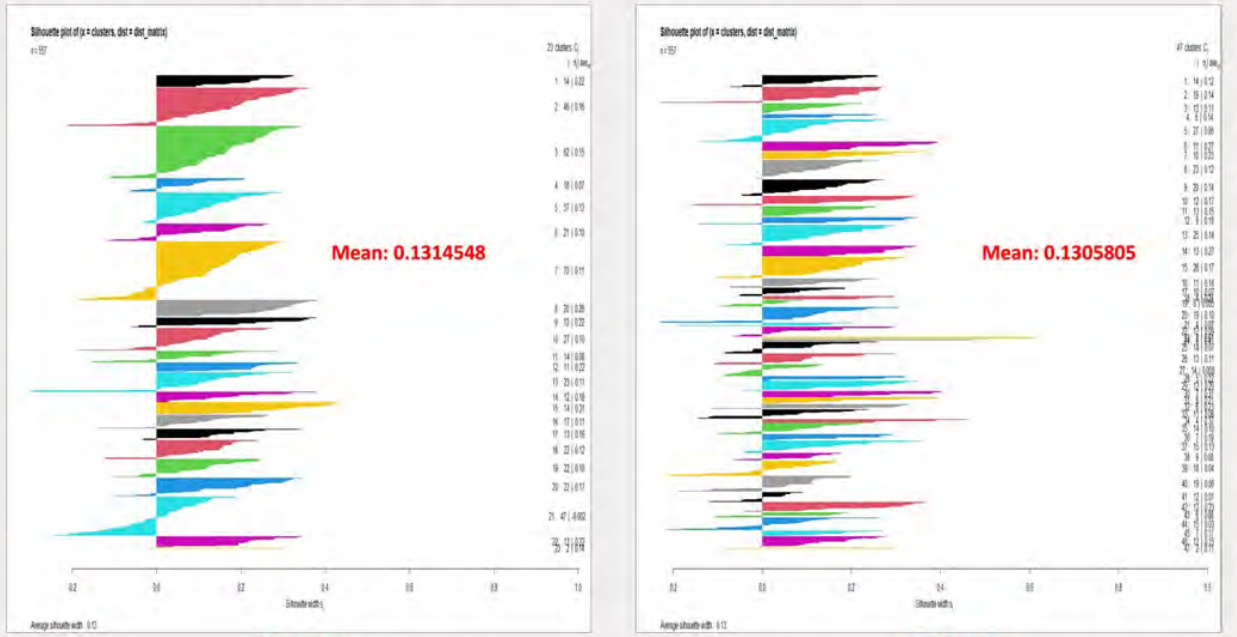
Knowledge-clusters Preliminary Results:

TABLE 3. Silhouette Score Table of Knowledge-levels Squared (Job Zones 3, 4 and 5)

Number of knowledge clusters	Mean Silhouette Score
15	0.1189706
17	0.1174186
18	0.12220657
19	0.1254772
20	0.1286605
23	0.1314548
24	0.1264017
25	0.1280238
30	0.1220087
35	0.1291828
40	0.1270338
45	0.1300078
47	0.1305805

Source: Processed by PCRD.

FIGURE 8. Silhouette Plots of Knowledge-levels Squared



Source: Based on Average Silhouette Method. Processed by PCRD.

Silhouette plots revealed that 23 clusters are feasible. A few clusters had small number of occupations, hence merged with larger clusters. Qualitative assessment of individual clusters. Final 21 knowledge-based occupation clusters for U.S.

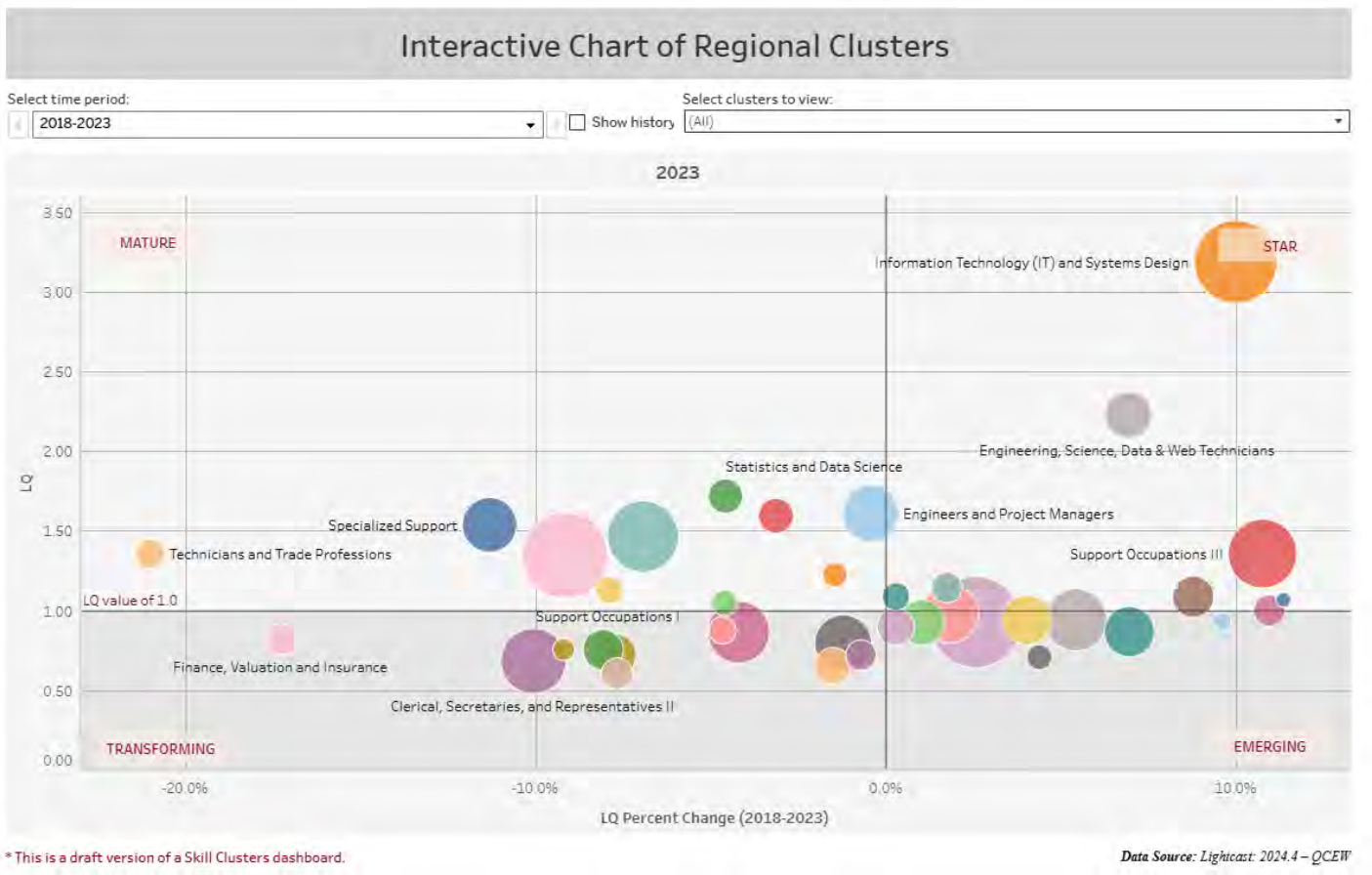
TABLE 4. 21 Knowledge-based Occupation Clusters

Knowledge Cluster #	Knowledge Occupation Cluster Name	Number of Occupations	Job Zones
C1	C1: Finance, Economics, and Accounting Professions	16	3,4,5
C2	C2: Healthcare and Medical Science Practitioners and Specialists	46	3,4,5
C3	C3: Healthcare and Medical Science Technicians	40	3,4,5
C4	C4: Legal, Environmental Compliance and Enforcement Professions	14	3,4,5
C5	C5: Business Management and Client Services Professions	31	3,4
C6	C6: Education, Religion, History, and Cultural Professions	18	3,4,5
C7	C7: Engineering and Technical Trade Technicians	65	3,4,5
C8	C8: Engineering and Advanced Technology Professions	31	3,4,5
C9	C9: Agriculture, Food and Natural Sciences Professions	21	3,4,5
C10	C10: Environmental and Earth/Geo Science Professions	18	3,4,5
C11	C11: Transportation and Logistics Professions	16	3,4,5
C12	C12: Humanities, Sciences and Engineering Post Secondary Education and Knowledge Creation	40	3,4,5
C13	C13: Architecture, Landscape and Energy Professions	13	3,4,5
C14	C14: Arts, Media, Communications and Creative Professions	33	3,4,5
C15	C15: Mental Health and Social Service Professions	14	4,5
C16	C16: Microbiological, Physical, Mathematical, and Statistical Sciences Professions	20	4,5
C17	C17: Office, Publishing, Information Management and Curation Support Professions	16	3,4,5
C18	C18: Operations, Production and Safety Management Professions	16	3,4
C19	C19: Information Technology and Telecommunication Professions	22	3,4,5
C20	C20: Vocation, Management, Community and Childcare, and Personal Care Professions	56	3,4,5
C21	C21: Public Safety and Law Enforcement Professions	11	3,4

Source: Processed by PCRD.

Knowledge-clusters

Cluster data dashboards are in development for the 400+ Economic Development Districts in the U.S.



Contributors

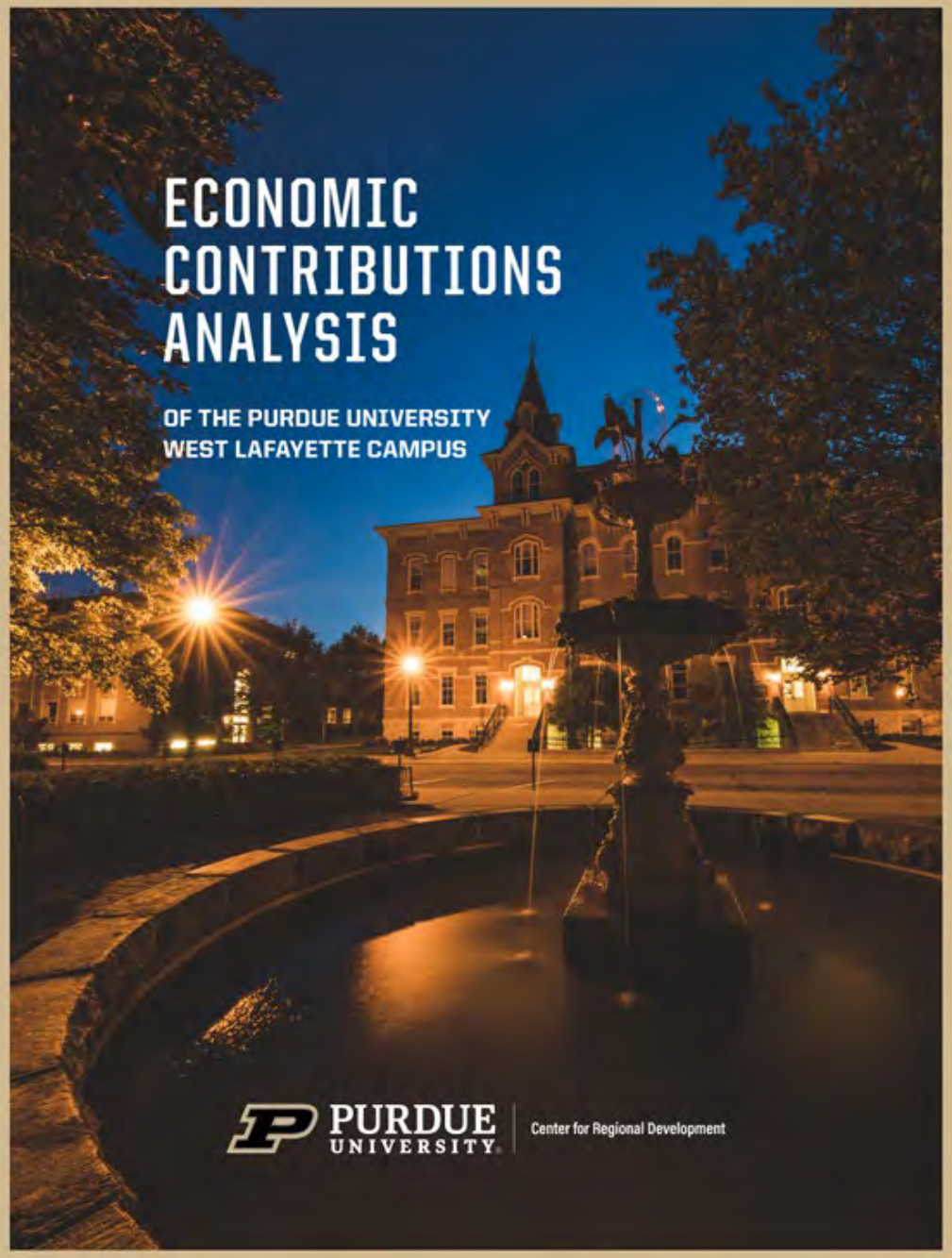
Research Team

Indraneel Kumar, Ph.D.
Annie Cruz-Porter, Ph.D.
Andrey Zhalnin, Ph.D.
Melinda Grismer, Ph.D.
Mark White, Ph.D.
Christine Nolan
Roberto Gallardo, Ph.D.
Hyewon Shin
Rachel Zhang
Ben St. Germain
Bertin Mbongo
Karen Ivanna Siller
Ryan Maluchnik
Tyler Wright
Taoyi Sun
Cheyanne Geideman

Data Analysis

Indraneel Kumar, Ph.D.
Taoyi Sun
Andrey Zhalnin, Ph.D.
Ben St. Germain
Bertin Mbongo

Disclaimer: This presentation is prepared by Purdue Center for Regional Development using Federal funds under award #ED23RNA0G0174 from the Economic Development Administration, U.S. Department of Commerce. The statements, findings, conclusions, and recommendations are those of the author(s) and do not necessarily reflect the views of Economic Development Administration or the U.S. Department of Commerce.



REMI User Conference 2025 Ann Arbor, MI

Economic Contributions and Impacts Analysis
Purdue University West Lafayette Campus 2024

Presented by: Indraneel Kumar, Ph.D., AICP

Introduction

- There is a need for colleges and universities to convey their economic value to the decisionmakers.
- The present study uses 2024 calendar year direct expenditure categories:
  - O&M (Operating expenses and maintenance include R&D expenditures): $2.3 Billion
  - New construction and related expenditures (include consultants): $65 Million
  - Off campus student expenditures: $785 Million
  - Spending by visitors to athletic events who were not students: $40 Million
- O&M based on the seven Integrated Postsecondary Education Data System (IPEDS) categories.
  - Instruction, scholarship and departmental research; academic support and services; operation, repair and maintenance of plants; office administrative services; auxiliary enterprises; public services; scientific R&D
- The study is based on expenditures. Revenues and monetized values of inventions and discoveries are not included.
- Similarly, Purdue Research Foundation (PRF) Research Park, Purdue Applied Research Institute (PARI) and multiple manufacturing establishments based on industry-university partnerships are not included.
- Local economic impacts of attendees to conferences, and visitors to convocations are not included.
- Increment in professional lifetime earnings of Purdue graduates are not included. Higher earnings by STEM disciplines than non-STEM disciplines.

Literature Review 1

- Impact analysis of universities requires a multifaceted approach unlike impact analysis of public infrastructure.
- The land grant universities have distinct mission for teaching, research and discovery, and public engagement through Agricultural and Engineering Extension services providing services to communities and regions.
- Most of the colleges and universities are operating since more than hundred years. Identifying new impacts to the local and regional economies from university operations can be challenging. Establishing a counterfactual or "how much the area is better or worse off with or without a university" (Siegfried, Sanderson and McHenry 2007).
- A few researchers suggested to incorporate stochastic methods (Monte Carlo Simulation) especially for expenditure areas such as student spending to augment the present methodology (Pastor, Perez and Guevara 2013).
- General methodological approach suggested is: Supply side (university developing human capital, knowledge creation, etc.) vs. Demand side (university as a consumer spending on various goods and services).
- Economic impacts can be generative (new economic growth), distributive (relocating jobs and workers), and transfer (intergovernmental transactions between federal and state and local governments) (CS, Cervero and Aschauer 1998).
- Universities can create substantive generative impacts from external R&D dollars, new constriction of labs and buildings, spending by out of town visitors, etc.).

Literature Review 2

- Generally, we know that economic impacts include assessment of direct, indirect and induced impacts.
- Such and linear and additive framework conceptually might not work for universities engaged in developing "new knowledge" and "new skilled human capital".
- Computable/Applied General Equilibrium (CGE) with forecasting capacity is better than the Input-Output (IO).
- Potential increase in lifelong earnings of graduates remaining in the state happens in future years.
- There are questions with sporting events as well! Some researchers have found that considering "net" instead of "gross" economic activities resulted into modest impacts from sports events (Baade, Baumann and Matheson 2011).
- Some cautionary tips from the literature:
  - Appropriate delineation of local and regional geography for impact assessment (Siegfried, Sanderson and McHenry 2007; Pastor, Perez and Guevara 2013; Quddus et al. 2022)
  - Avoid double counting to ensure distribute spending in mutually exclusive categories (Siegfried, Sanderson and McHenry 2007)
  - Focus on net economic activities (Baade, Baumann and Matheson 2011)
- The missing piece in literature is on considering the new trend of onsite, hybrid and remote workers.

Input Data 2024

- $2.3 Billion contains R&D expenditures of $394.5 Million.

$2.3B
operations, maintenance, and R&D

$785M
off campus student expenditures

$65M
new construction

$40M
athletic event visitor expenditures

- Fall 2024 Statewide Enrollment was 74,800
- Fall 2024 West Lafayette Campus Enrollment was 58,000
- Fall 2024 West Lafayette Purdue Employees was 19,027
- Tuition is frozen for 14 consecutive years

Data Collection and Model Details 1

- $785 million off campus student expenditures are based on student spending surveys conducted by the university:
  - Spending on housing, food, books, transportation and miscellaneous activities were estimated for off campus
  - Spending on books, transportation and miscellaneous activities were estimated for on campus.
- Miscellaneous spending was distributed into various consumer spending categories based on assumptions.
- We knew ticket purchases by students and non-students in 2024:
  - Placer.ai was used to geofence Ross-Ade Stadium and all visitors spending 30-minutes or more were filtered
  - Placer.ai was used to collect visitors data for select high profile and general games during 2024
  - Placer.ai provides originating zip codes for visitors
  - The average proportions of visitors were estimated for Tippecanoe County, rest of Indiana, and rest of USA
  - The spending on food, fuel, and accommodations were estimated based on assumptions
- College amateur sports and athletic events can stimulate local economies if organized at regular intervals (Brewer and Freeman 2015). Extensive visitor surveys are needed to estimate the dollar expenditures and categories. Assumptions were made due to the lack of survey.
- 89.2% of employees live in Tippecanoe County, 7.3% in the rest of Indiana, and 3.5% lived outside of Indiana.

Data Collection and Model Details 2

1) REMI Policy Insight + v. 3.2.1 is used. It is a 3-region model: Tippecanoe County, rest of Indiana, and rest of USA.
2) Identifying the appropriate geography is the key. For example, Purdue Extension office is in every county and services include technical assistance in agriculture, and connecting to main campus resources such as Local Technical Assistance Program for civil engineering. Manufacturing Extension Partnership has several offices in the state and provide statewide services to support industries.
3) The focus was also identifying specific categories of spending where induced investments from expenditures are nullified, such as operations and maintenance, salaries, etc.
4) R&D investments are not nullified because the research induces new ventures in Purdue Research Park and new collaboration with industries creating new business establishments and jobs in Indiana.
5) The department-level information on alumni occupations and job locations could not be included in the model because of incomplete information. The objective is to include this information in the next iteration.
6) Graduates from engineering, pharmacy, management, and veterinary science schools earn much higher than graduates from arts and letters.

Results

2024 CONTRIBUTIONS

$6.5

BILLION

to county, state, and U.S. GDP

GENERATED

28,900

JOBS IN INDIANA

+19,000 jobs at Purdue WL

$3B+$1B

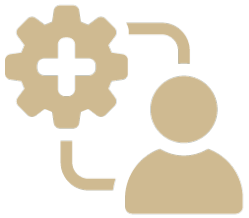
county's GDP + state's GDP

$2.45B

nation's GDP

13,000

new state residents



+2.5%

labor force participation in Tippecanoe County

Results 1: Economic indicators (Differences)

Category	Units	Tippecanoe 2024	Rest of Indiana 2024	Total Indiana 2024
Total Employment	Thousands (Jobs)	39.07	8.88	47.95
Residence Adjusted Employment	Thousands	32.57	15.02	47.59

Category	Units	Tippecanoe 2024	Rest of Indiana 2024	Total Indiana 2024
Gross Regional Product	Billions of Fixed (2024) Dollars	3.02	1.02	4.04
Output	Billions of Fixed (2024) Dollars	4.91	1.75	6.66
Value-Added	Billions of Fixed (2024) Dollars	3.02	1.02	4.04
Personal Income	Billions of Fixed (2024) Dollars	1.40	0.93	2.33
Disposable Personal Income	Billions of Fixed (2024) Dollars	1.22	0.81	2.04

Source: REMI simulations run by PCRD

Results 2: GDP and PCI (Differences)

Table 2: Gross Regional/Domestic Product (Differences) 2024

Region	Units	2024
Tippecanoe County, IN	Billions of Fixed (2024) Dollars	3.02
Rest of Indiana	Billions of Fixed (2024) Dollars	1.02
Rest of U.S.	Billions of Fixed (2024) Dollars	2.45
Total	**Billions of Fixed (2024) Dollars**	**6.49**

Source: REMI simulations run by PCRD

Table 3: Disposable Personal Income Per Capita (Differences) 2024

Category	Units	Tippecanoe 2024
Real Disposable Personal Income	Billions of Fixed (2024) Dollars	1.22
Real Disposable Personal Income per Capita	Thousands of Fixed (2024) Dollars	4.08

Source: REMI simulations run by PCRD

Results 3: Demographic indicators (Differences)

Table 4: Labor Force Components (Differences) 2024

Category	Units	Tippecanoe 2024
Labor Participation Rate	Rate	0.025
Labor Force	Thousands	8.12

Source: REMI simulations run by PCRD

Table 5: Population (Differences) 2024

Region	Units	Year 2024
Tippecanoe County, IN	Thousands	8.6
Rest of Indiana	Thousands	4.2
Total Indiana	**Thousands**	**12.8**

Source: REMI simulations run by PCRD

Discussions and Next Steps

1) Economic contributions and impacts are not limited to only Tippecanoe County, but spillover to the rest of Indiana and the rest of USA.
2) There are positive demographic contributions and impacts in terms of economic migrants and movers increasing and retaining the population as well as increase in the labor force participation rate.
3) $3.19 B of expenditures in select categories created $6.5 B of total changes in GDP including Tippecanoe County, rest of Indiana, and the rest of USA.
4) There is a significant contribution to the real disposable personal income and the real disposable personal per capita income increases by almost $4K in Tippecanoe County in 2024.
5) Prepare guidelines for data collection for future iterations of the economic contributions and impacts study.
6) Experiment with incorporating stochastic methods for data collection.

Report and Acknowledgement

https://pcrd.purdue.edu/purdue-economic-impact

Report Author

Indraneel Kumar, Ph.D.

Reviewers

Roberto Gallardo, PhD
Alejandro Madrigal (REMI)
Chris Judson (REMI)
Peter Evangelakis, PhD (REMI)

Team Members

Office of Engagement
Roberto Gallardo, PhD

Purdue Center for Regional Development (PCRD)
Indraneel Kumar, PhD
Ben St. Germain

Institutional Data Analytics + Assessment (IDA+A)
Laura Beth Todd
Andrea Pluckebaum
Michael S. Wakolbinger

University of Georgia (Michael Adjemian, Ph.D.) and University of Connecticut (Fred Carstensen, Ph.D., Peter Gunther, and Theophilus Djaba, Ph.D.) faculty members for interviews and insights.

PURDUE UNIVERSITY | Center for Regional Development

Contact Us

1341 Northwestern Avenue
Purdue Schowe House
West Lafayette, IN 47906

765-494-7273

pcrd@purdue.edu

Purdue Center for Regional Development

seeks to pioneer new ideas and strategies that contribute to regional collaboration, innovation and prosperity.

www.pcrd.purdue.edu

Schowe House photography is by PCRD staff.