

## Finding Opportunities to Diversify Regional Economies with Lessons for Policy from a Case Study on Hawai‘i

Steven Bond-Smith<sup>1\*</sup>, Sumit Ilamkar<sup>2</sup>

1. University of Hawai‘i  
Economic Research  
Organization (UHERO),  
University of Hawai‘i at  
Mānoa, Honolulu, Hawai‘i,  
United States

2. Department of Economics,  
University of Hawai‘i at  
Mānoa, Honolulu, Hawai‘i,  
United States

\*Corresponding author.  
Email: [bondsmi@hawaii.edu](mailto:bondsmi@hawaii.edu)

Many regions aim to diversify their economies. But objective tools are rarely used to find diversification opportunities. We propose using the principle of relatedness to identify potential opportunities. Specifically, we use objective measures to identify small or new industries that fit well with the region's existing strengths. We summarize a case study on diversifying Hawai‘i's economy and discuss policy implications. Even with this clear method to identify opportunities, there will be barriers to developing these industries. New industries will therefore need a targeted policy response that addresses barriers to development.

### Keywords

Diversification, Regional Development, Resilience, Relatedness

### JEL Classification

O18, O25, R10, R11

**Submitted:** 08/15/2024 **Accepted:** 07/10/2025

### Acknowledgements

We are grateful for many helpful comments and suggestions from the two editors, John Winters and Heather Stephens, and two anonymous referees. We are also grateful for discussions with seminar participants at the Hawaii Economic Association; State of Hawai‘i Department of Business, Economic Development, and Tourism; Maui County Council Agriculture, Diversification, Environment, and Public Transport Committee; Maui Economic Recovery Commission; Climate-Resilient Food Innovation Network Workshop and Maui Economic Development Board Ka Ipu Kukui Fellows and the University of Hawai‘i Economic Research Organization Pau Hana. The original outreach report was partly funded by the Climate-Resilient Food Innovation Network through the NSF Engines Development Award 2305455.

© 2025 Bond-Smith and Ilamkar. CC-BY. This is an open access article published under a CC-BY 4.0 license.

Bond-Smith, S., & Ilamkar, S. (2025). Identifying opportunities to diversify regional economies with lessons for policy from a case study on Hawai‘i. *Reaching Regions*, 1, 18310.  
<https://doi.org/10.31274/rreg.18310>

## 1. Introduction

Many regions of the United States specialize in single sectors, including the Rust Belt, Corn Belt, and California's Central Valley, as well as tourist areas like Hawai'i that serve in hospitality and entertainment. Specializing provides gains by trading for goods that would be more costly for a region to produce themselves. Increasing an industry's scale can also boost its productivity, reinforcing specialization. Yet specializing also makes a region more vulnerable to both shocks and long-term trends and may prevent other industries from emerging. Over-specialized regions will not achieve their potential, and are unprepared to take advantage of growth opportunities.

In response, the Economic Development Administration promotes strategies for resilience that include diversifying a region's industrial base. As a result, many regions target diversification in their Comprehensive Economic Development Strategies. However, regions will be more likely to successfully diversify if they use objective tools to identify the most promising local opportunities.

In this article, we use relatedness to identify potential opportunities to diversify, using Hawai'i as a case study. The principle of relatedness (Hidalgo et al., 2018) explains that economies grow by branching out from existing industries that use capabilities similar to their existing strengths (Hidalgo et al., 2007; Neffke et al., 2011). Related industries need similar inputs and conditions to thrive (Hausmann & Hidalgo, 2011), and adding a related industry only requires a little new "know-how" for the region to have all the necessary factors for that industry (Hausmann, 2016). We use our case study to also explore general implications for regional economic diversification strategies.

## 2. Methods and Data

Regional economists and economic geographers now promote objective tools for regional development based on the principle of relatedness (Balland et al., 2019).<sup>1</sup> We adapt these tools to analyze opportunities for diversifying a region's industrial base. Specifically, we target related industries that are currently underperforming. See the appendix for details on specific calculations and data limitations.

### 2.1 Data

To measure economic activity, we use US Census Bureau County Business Patterns (CBP) employee count data from 2021 for all counties in the United States. This data specifies employment using six-digit North American Industry Classification System (NAICS) industries. Since CBP excludes agriculture and aquaculture, we supplement CBP data with location quotients (LQ) for agriculture NAICS codes from the Quarterly Census of Employment and Wages (QCEW) for March 2021 (the same month as the CBP data).

CBP data is highly accurate; however, data is suppressed if there are fewer than three firms present, and noise is infused in the firm-level data. Similarly, the QCEW suppresses data in more industries than does CBP, so we limit its use only to industries missing from CBP. We do

---

<sup>1</sup> See Boschma (2017) for a summary of research on relatedness. Variations of this approach were used to support economic development in Western Australia (Bond-Smith et al., 2019; Hausmann et al., 2021), South Africa (Hausmann et al., 2023), and Sri Lanka (Malaloda et al., 2018), among others.

not expect that data suppression or noise biases our results, though they may affect accuracy if there are too few counties hosting a six-digit industry. Therefore, we exclude industries that are hosted in less than 20 counties nationally. We include all US counties, including rural areas, since relatedness is not only defined by the co-location of industries in urban areas.

Data suppression in the CBP data can be an issue for calculating LQs in industries with fewer than three firms, which usually only affects smaller counties. Fortunately, large employers with few competitors in small counties should be easy to identify, allowing for manual corrections of falsely identified opportunities.<sup>2</sup> This limitation is not a major concern for our case study, as Hawai'i's counties are comparatively large and consultations with state and local governments, government departments, and economic development agencies have not revealed any issues with our conclusions.

## ***2.2 Location quotients of industries and clusters***

We first calculate LQs, which measure the local relative size of an industry, for NAICS six-digit industries for every US county. An LQ below 1 indicates the industry has a smaller share of the local economy than the national average. An LQ of 1 implies that the industry has the same share of the local economy as the national average. An LQ above 1 implies greater specialization in that industry.

## ***2.3 Measuring relatedness densities of industries and clusters***

Using LQs, we measure relatedness between two industries by estimating the probability that one industry has an LQ greater than 1, given that the other industry does as well.<sup>3</sup> This measure captures the tendency for specific pairs of industries to locate together, but is neutral to the reasons why they appear together. We then measure the intensity of relatedness between each industry and the portfolio of all industries in each county. This relatedness density metric for each industry and county shows the extent to which related industries are present in the county. Finally, we also calculate the density of related industries for specific industry clusters using the clusters defined in Delgado et al. (2016). Industry clusters are groups of industries (and the firms in them) that benefit from close proximity to each other. These benefits come from factors such as labor market pooling, shared knowledge, and supply-chain linkages. Beginning with Porter (2000), clusters have been a key tool for targeted economic development policies.

## ***2.4 Potential diversification opportunities***

We next use our measures of relatedness to find diversification opportunities. An industry with a high relatedness density suggests that the county is more likely to have the conditions needed for the industry to thrive. Many highly related industries may already be thriving. Thus, we look at related industries with LQs of less than 1, which indicates a smaller industry concentration and where growth would assist with economic diversification. However, if these industries are small,

---

<sup>2</sup> For example, a factory town would have only one employer in the factory's industry. With fewer than three firms it would not be included in the data, but it is likely to have high relatedness since its suppliers would be located nearby. Therefore, the industry would appear through our algorithm as a diversification opportunity with a low or zero LQ and high relatedness density, but it would be obvious to filter out of the results.

<sup>3</sup> Most calculations are performed using the EconGeo Package in R (Balland, 2017). The appendix describes technical details on the specific calculations.

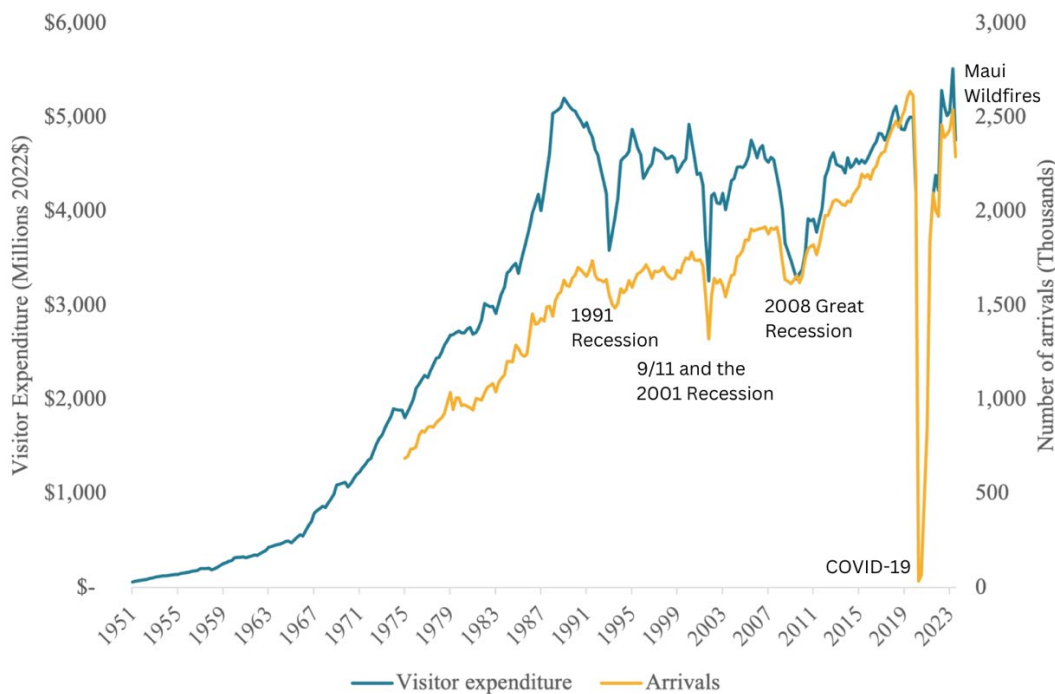
they may face significant barriers to expansion. Therefore, we refer to these industries as potential diversification opportunities. In addition, we use relatedness by cluster to identify which clusters fit well with the region's industrial structure. This is helpful to prioritize clusters that include both existing strengths and diversification opportunities.

Lastly, we exclude some industries from our conclusions as relatedness density is an imperfect proxy for feasibility. Our goal is to diversify by expanding underperforming industries; however, many such industries are infeasible if a critical natural resource is not available.

### 3. Potential Opportunities to Diversify Hawai'i's Economy

#### 3.1 The economic case for diversifying Hawai'i's economy

Hawai'i's economy is highly dependent on tourism and thus is exposed to short-term fluctuations in the US and global economies that affect this sector (Bond-Smith & Fuleky, 2023). Since 1990, tourist numbers have been regularly hit by shocks, including the 1991 recession, the 2001 recession (and 9/11), the Great Recession, COVID-19, and most recently, the Maui wildfires (see Figure 1). Hawai'i's dependence on tourism also makes it vulnerable to long-term trends. Tourism grew a lot in the 30 years after statehood in 1959. But tourism spending has stagnated in the next 30 years (Bond-Smith, 2024).

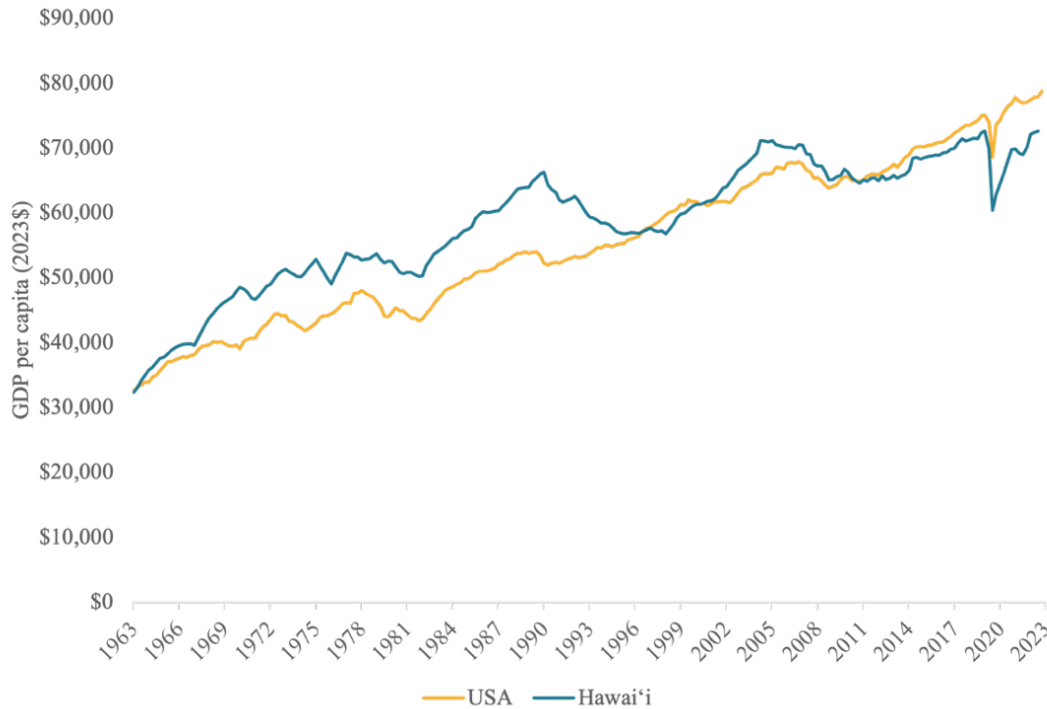


**Figure 1. Quarterly real visitor spending (2023 \$) Q1 1951 to Q1 2024 and quarterly tourist arrivals Q1 1975 to Q1 2024.**

*Data source:* University of Hawai'i Economic Research Organization (UHERO) (n.d.).

*Notes:* Visitor spending deflated to 2023 dollars using the Honolulu CPI. Both series are seasonally adjusted by UHERO.

As a result, Hawai‘i’s economic growth has lagged since the Great Recession. Arguably, it has lagged for even longer, since the so-called *lost decade* in Japan in the 1990s. Hawai‘i’s per capita GDP is now around 8% lower than US per capita GDP (see Figure 2) and the gap is forecast to continue widening. Furthermore, Hawai‘i’s high cost of living means that GDP per capita in price parity terms is now about 17% lower than the rest of the United States. This performance is similar to "left-behind places" (Hendrickson et al., 2018).



**Figure 2. Quarterly real per capita GDP, United States and Hawai‘i, 1964–2023.**

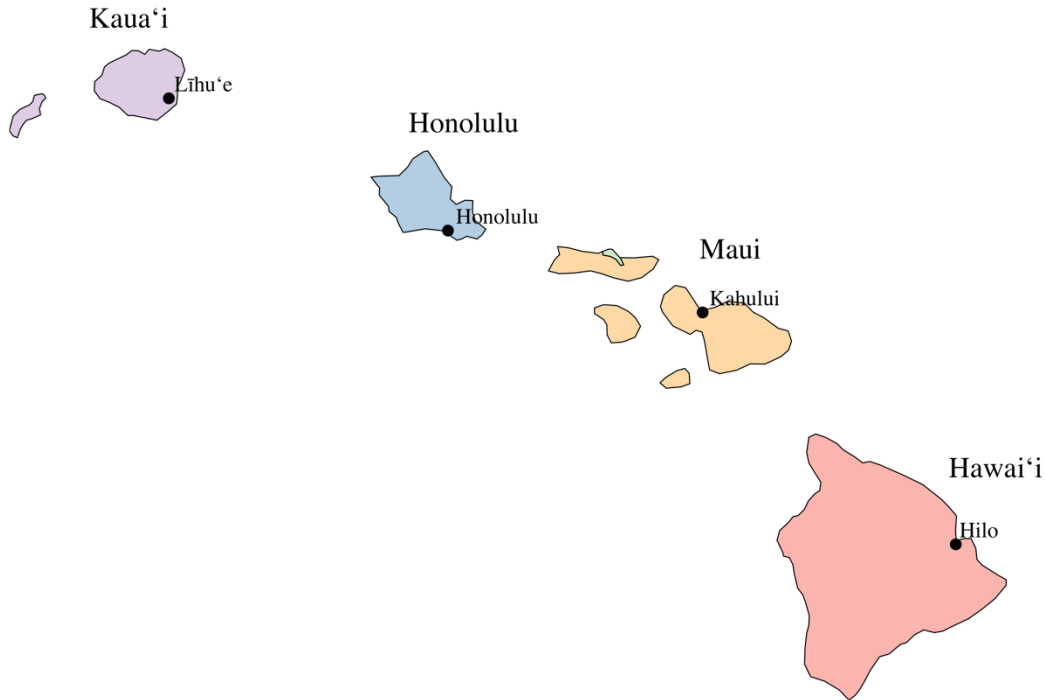
*Data source:* Authors’ calculations with UHERO (n.d.) data.

*Notes:* US per capita GDP is deflated using the US GDP deflator. Hawai‘i per capita GDP is deflated using the Honolulu CPI. Honolulu CPI is seasonally adjusted by UHERO. US and Hawai‘i GDP per capita data are seasonally adjusted by the Bureau of Economic Analysis.

### 3.2 Opportunities for diversification

Using the methods described above, we identify diversification opportunities in all four major counties in Hawai‘i: Honolulu (O‘ahu), Maui, Hawai‘i Island, and Kaua‘i (See map in Figure 3). Figure 4 shows relatedness densities and LQs for Maui as an example, since Maui is more specialized in tourism than the other counties.<sup>4</sup> The size of the bubble indicates the number of workers in that industry. Industries further to the right are likely to be more feasible, based on our estimates of relatedness densities. A region’s current strengths (industries with higher shares of employment than the national average) are in the upper half of the chart, with LQs above 1. Diversification opportunities appear in the lower-right corner of the chart. These industries have LQs below 1, indicating they are currently small, and yet, they have higher relatedness densities, suggesting they are more viable.

<sup>4</sup> See Bond-Smith and Ilamkar (2024) for results and charts for counties not included here.



**Figure 3. Hawai'i's four main counties.**

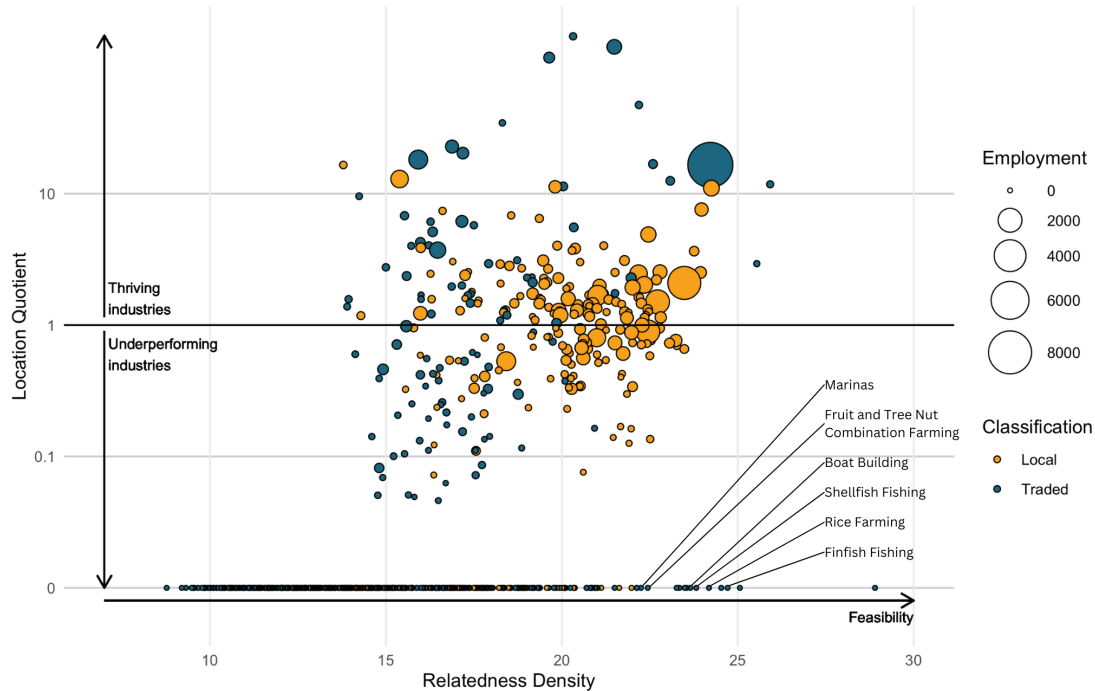
*Source:* USAboundaries R package (Mullen, 2022).

*Notes:* For data collection, the fifth county, Kalawao, is treated as part of Maui County.

In all four of Hawai'i's counties, several ocean-based industries—Finfish Fishing, Shellfish Fishing, Boat Building, Port and Harbor Operations, and Seafood Preparation and Packaging—stand out as potential diversification opportunities. In Honolulu, Shellfish Fishing and Seafood Product Preparation and Packaging will likely build on existing strengths in Finfish Fishing and Fish and Seafood Merchant Wholesalers. In all four counties, Marinas and Boat Building likely build on existing strengths such as Scenic and Sightseeing Transportation, Water. We are cautious of drawing conclusions for agriculture, but the results suggest that the opportunities are in niche tropical crops.<sup>5</sup>

Of the four counties in Hawai'i, Kaua'i has the fewest options for diversifying—residents might also consider hybrid-remote work in Honolulu, which offers a more diversified job market. Hawai'i Island has many more options to diversify its economy than the other three counties and its diversity creates more branching opportunities. However, unlike O'ahu, many of Hawai'i Island's related industries are small.

<sup>5</sup> Due to the limitations of the QCEW data and the broad categories that include tropical crops in NAICS classifications, a more detailed study of specific crops would be required to fully understand the possibilities.



**Figure 4. Identifying diversification opportunities in Maui County.**

*Notes:* Diversification opportunities are those with a higher relatedness density that are currently small. Some agriculture and fishing industries have an LQ of 0 to protect confidentiality in the source data. If these LQs are actually already greater than 1, they may not be suitable to expand. A few local and traded classifications may switch in the Hawai‘i context due to its island geography and specialization in tourism.

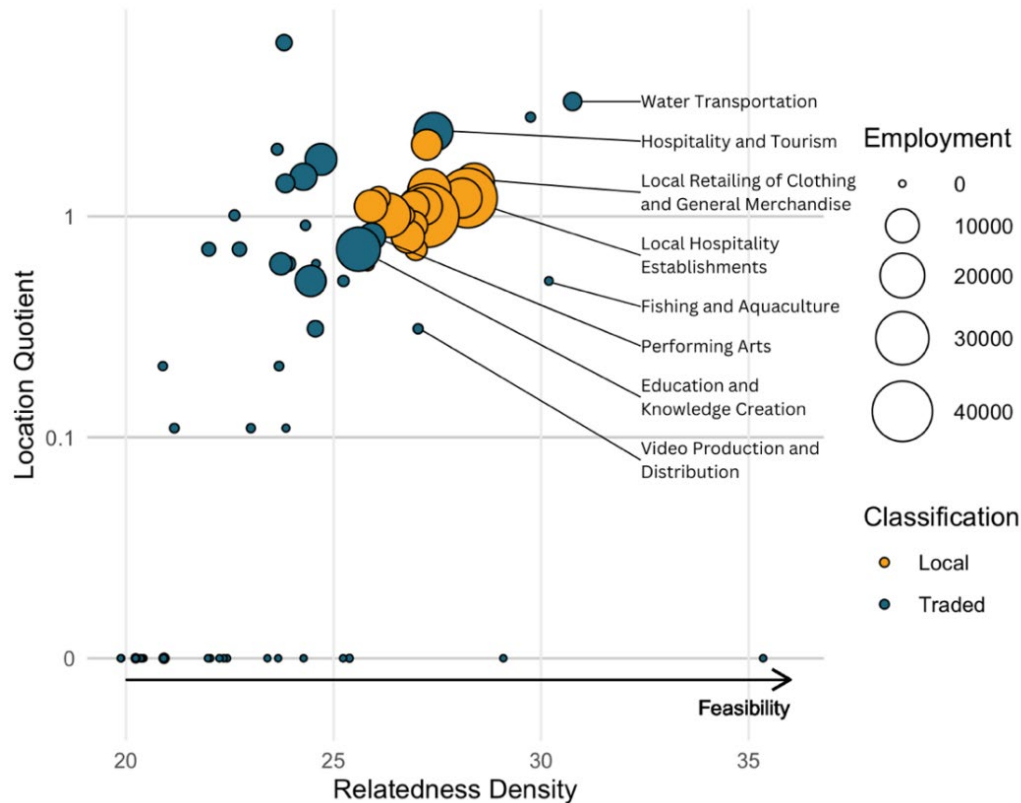
### 3.3 Clusters of strengths and opportunities

Prioritizing clusters based on relatedness density could also support diversification. Many regions currently prioritize clusters and industries that are already growing; however, growing clusters do not necessarily need more support. Relatedness highlights clusters with the potential to also support diversification.<sup>6</sup> Figure 5 shows relatedness densities by cluster for Honolulu as an example since it also hosts the strong clusters that appear in the other counties. Again, some clusters are disregarded if they are clearly not feasible.<sup>7</sup>

As expected, clusters serving tourists feature strongly in all four counties. The primary cluster serving tourists is Hospitality and Tourism; however, Local Hospitality Establishments, Local Retailing of Clothing and General Merchandise, Jewelry and Precious Metals, and Performing Arts also serve tourists, despite being classified as local—in the Hawai‘i context, these can be considered as traded. The Fishing and Aquaculture cluster stands out with a higher relatedness density across all four counties. This result suggests that several industries within the cluster could expand. However, regulation must balance ecological concerns and commercial viability for this cluster to thrive.

<sup>6</sup> Many of Hawai‘i’s so-called local clusters also serve tourists and some traded clusters are not traded between islands, thus, we occasionally treat the local and traded classifications differently in this study. For example, Delgado et. al. (2016) classify Full-Service Restaurants as local, but this industry is large in Hawai‘i because it serves tourists, so we treat it as traded in our context.

<sup>7</sup> For example, Coal Mining is a potential cluster for all islands but this is clearly not feasible without coal deposits and is disregarded.



**Figure 5. Prioritizing clusters in Honolulu County.**

*Notes:* Clusters with higher relatedness densities include more existing strengths and diversification opportunities. Some local and traded classifications can be reversed in the Hawai‘i context due to its island geography and specialization.

O‘ahu hosts an Education and Knowledge Creation cluster. However, its strength is underestimated since the data excludes public employment (i.e., employees of the University of Hawai‘i). The strongest diversification opportunity within the cluster is Research and Development in the Physical, Engineering, and Life Sciences. Yet this already features an LQ of around 1 on Maui and 1.2 on Hawai‘i Island. So, from a diversification perspective, this may also be a good cluster for the other counties too.

Video Production and Distribution features moderately in Honolulu, is also a moderate option for Maui, and has potential on Hawai‘i Island and Kaua‘i. Hawai‘i has hosted many film and television productions, but its employment level is still relatively low. The starting point is to determine the missing capabilities that may be limiting its growth. Previous research notes that the annual tax credit cap on all productions may be too low, disincentivizing an expanded industry (La Croix & Mak, 2021).

The Water Transportation cluster features in results for both Honolulu and Maui County. This cluster is predominantly about freight, including Marine Cargo Handling and Coastal Freight Transportation industries; however, it also includes the diversification options such as Boat Building and Deep Sea Freight Transportation. Though hosting the latter industry on O‘ahu may require policy reforms, which seems unlikely.

## 4. Conclusions for Policy

This article proposes a new approach in US regional policy: addressing issues in underperforming but potentially promising industries prioritized using objective measures, illustrated with Hawai‘i as a case study. The main limitation is its backward-looking focus does not observe emerging industries without a US presence or NAICS code. Still, for many lagging regions, aspiring to host entirely new industries is unrealistic. Policies can nonetheless still address barriers to new industries to enable market discovery. Our approach is not a blanket solution but shows how relatedness offers objective evidence to identify opportunities.

Policy design is key. Industries identified as diversification opportunities have not emerged naturally due to barriers or bottlenecks—likely market or government failures. Counties may have the necessary capabilities, but various failures or developmental bottlenecks mean that new industries are blocked. While economists are wary of “picking winners”—a process relying on public information and vulnerable to bias—our approach avoids this. Relatedness reveals where the market implies success is possible but unrealized, suggesting a barrier to development.

Effective policy design uncovers these barriers and addresses them cost-effectively, often requiring deep industry knowledge. On this basis, diversification policies would invite expert proposals to address barriers to growth in potential diversification opportunities. Relatedness helps to prioritize policy efforts, not by picking winners, but by identifying where barriers impede likely success. As industries grow, local relatedness changes, creating further opportunities. Relatedness thus supports long-term diversification by starting with the most feasible targets. Alshamsi et. al. (2018) also suggest extending the algorithm to include targeting highly connected activities that generate more diversification opportunities in the future.

Overall, our study leads to two overarching conclusions for policy. First, objective tools like relatedness help identify viable industries and can be used to allocate resources strategically. But secondly, diversification is hard. Highly-related, underperforming industries face substantial barriers. If these are insurmountable, the economy may be specialized out of necessity, and not over-specialized. Even then, identifying and addressing bottlenecks remains worthwhile.

## References

- Alshamsi, A., Pinheiro, F. L., & Hidalgo, C. A. (2018). Optimal diversification strategies in the networks of related products and of related research areas. *Nature Communications*, *9*, 1328. <https://doi.org/10.1038/s41467-018-03740-9>
- Balland, P-A. (2017). Economic geography in R: Introduction to the EconGeo package. Papers in Evolutionary Economic Geography (PEEG) 1709. Utrecht University, Department of Human Geography and Spatial Planning. <https://peeg.wordpress.com/2017/05/09/17-09-economic-geography-in-r-introduction-to-the-econgeo-package/>
- Balland, P-A., Boschma, R., Crespo, J., & Rigby, D. L. (2019). Smart specialization policy in the European Union: Relatedness, knowledge complexity and regional diversification. *Regional Studies*, *53*(9), 1252–1268.
- Bond-Smith, S. (2024). Diversifying Hawai‘i’s specialized economy: A spatial economic perspective. *Economic Development Quarterly*, *38*(1), 40–59. <https://doi.org/10.1177/08912424231183941>
- Bond-Smith, S., Dockery, A. M., Duncan, A., Kiely, D., & Salazar, S. (2019). Future-proofing the WA Economy: A roadmap to industrial diversification and regional growth. Bankwest Curtin Economics Centre, BCEC Focus on Industry Series, No. 4. <https://bcec.edu.au/assets/2019/08/BCEC-Future-Proofing-the-WA-Economy-report.pdf>
- Bond-Smith, S., & Fuleky, P. (2023). The effects of the pandemic on the economy of Hawai‘i. *Current Issues in Tourism*, *26*(23), 3846–3852. <https://doi.org/10.1080/13683500.2022.2151875>
- Bond-Smith, S., & Ilamkar, S. (2024). Potential opportunities to diversify the economy of Hawai‘i. University of Hawai‘i Economic Research Organization (UHERO), UHERO Report, July 2. <https://uhero.hawaii.edu/wp-content/uploads/2024/07/PotentialOpportunitiesToDiversifyTheEconomyOfHawaii.pdf>
- Boschma R. (2017). Relatedness as driver of regional diversification: A research agenda. *Regional Studies*, *51*(3), 351–364. <https://doi.org/10.1080/00343404.2016.1254767>
- Delgado, M., Porter, M., & Stern, S. (2016). Defining clusters of related industries. *Journal of Economic Geography*, *16*(1), 1–38. <https://doi.org/10.1080/00343404.2016.1254767>
- Hausmann, R. (2016). Economic development and the accumulation of know-how. *Welsh Economic Review*, *24*, 13–16. <https://doi.org/10.18573/j.2016.10049>
- Hausmann, R., Barrios, D., Grisanti, A., Kasoolu, S., O’Brien, T., Protzer, E., Sanghvi, S., Taniparti, N., & Tapia, J. (2021). Western Australia research findings and policy recommendations. CID Faculty Working Paper No. 395, April. Center for International Development, Harvard University. <https://www.hks.harvard.edu/centers/cid/publications/faculty-working-papers/western-australia-policy-recommendations>
- Hausmann, R., & Hidalgo, C.A. (2011). The network structure of economic output. *Journal of Economic Growth*, *16*, 309–342. <https://doi.org/10.1007/s10887-011-9071-4>
- Hausmann, R., O’Brien, T., Fortunato, A., Lochmann, A., Shah, K., Venturi, L., Enciso-Valdivia, S., Vashkinskaya, E., Ahuja, K., Klinger, B., Sturzenegger, F., & Tokman, M. (2023). Growth through inclusion in South Africa. CID Faculty Working Paper No. 434, November. Center for International Development, Harvard University. <https://growthlab.hks.harvard.edu/sites/projects.iq.harvard.edu/files/growthlab/files/2023-11-cid-wp-434-south-africa-growth-through-inclusion.pdf>

- Hendrickson, C., Muro, M., and Galston, W. A. (2018). Countering the geography of discontent: Strategies for left-behind places. Metropolitan Policy Program at Brookings. Brookings. [https://www.brookings.edu/wp-content/uploads/2018/11/2018.11\\_Report\\_Countering-geography-of-discontent\\_Hendrickson-Muro-Galston.pdf](https://www.brookings.edu/wp-content/uploads/2018/11/2018.11_Report_Countering-geography-of-discontent_Hendrickson-Muro-Galston.pdf)
- Hidalgo, C. A., Balland, P.-A., Boschma, R., Delgado, M., Feldman, M., Glaeser, E., He, C., Kogler, D. F., Morrison, A., Neffke, F., Rigby, D., Stern, S., Zheng, S., and Zhu, S. (2018). The principle of relatedness. In: A. J. Morales, C. Gershenson, D. Braha, A. A. Manai, & Y. Bar-Yam (Eds.) *Unifying themes in complex systems IX*. Springer. [https://doi.org/10.1007/978-3-319-96661-8\\_46](https://doi.org/10.1007/978-3-319-96661-8_46)
- Hidalgo, C. A., Klinger, B., Barabási, A.-L., & Hausmann, R. (2007). The product space conditions the development of nations. *Science*, 317(5837), 482–487. <https://doi.org/10.1126/science.1144581>
- La Croix, S., and Mak, J. (2021). Understanding the role of the Hawaii film/TV/digital production tax credit in diversifying the Hawaii economy. UHERO Brief, March 15. <https://uhero.hawaii.edu/understanding-the-role-of-the-hawaii-film-tv-digital-production-tax-credit-in-diversifying-the-hawaii-economy/>
- Li, Y., & Neffke, F. M. H. (2024). Evaluating the principle of relatedness: Estimation, drivers and implications for policy. *Research Policy*, 53(3), 104952. <https://doi.org/10.1016/j.respol.2024.104952>
- Malalgoda, C., Samaraweera, P. & Stock, D. (2018). Targeting sectors for investment and export promotion in Sri Lanka. Growth Lab, Harvard Center for International Development. <http://www.tinyurl.com/27ndc9jn>
- Mullen, L. (2022). USAboundaries: Historical and contemporary boundaries of the United States. rOpenSci. Available at: <https://ropensci.org/packages/usaboundaries>
- Neffke, F., Henning, M. & Boschma, R. (2011). How do regions diversify over time? Industry relatedness and the development of new growth paths in regions. *Economic Geography*, 87, 237–265. <https://doi.org/10.1111/j.1944-8287.2011.01121.x>
- Porter, M. E. (2000). Location, competition, and economic development: Local clusters in a global economy. *Economic Development Quarterly*, 14(1), 15–34. <https://doi.org/10.1177/089124240001400105>
- University of Hawai‘i Economic Research Organization (UHERO). (n.d.). UHERO Data Portal. Retrieved 15 January 2025. <https://data.uhero.hawaii.edu/>

## Appendix

### *A.1 Data*

Our calculations require a large dataset of many industries and locations in order to accurately estimate statistical relationships. We use CBP employee count data from 2021 for all counties in all 50 US states. This annual series of subnational data includes employment by NAICS six-digit industries. Since the scope of CBP excludes agriculture and aquaculture, we supplement it with data from the QCEW in Quarter 1 Month 3 of 2021 (the same month as CBP data).

The CBP is our preferred data source. CBP data is highly accurate since it is extracted from the Business Register, a database of all known companies maintained by the US Census Bureau; however, there are some limitations. To protect confidentiality the CBP suppresses data if an industry has fewer than three businesses in a county. When employment in some industries is not reported it is effectively treated as zero when it is, in fact, not zero. The QCEW protects confidentiality by suppressing data in many more industries. This is one reason we prefer CBP, as data is suppressed in far fewer industries, so we only use the QCEW for agriculture, which is not reported in the CBP.

To further protect confidentiality in the CBP, noise is added to the employment count in industries with only a few businesses in a county. Distortions from this noise are reduced in our calculations by including more counties. We reduce this risk by including all US counties and excluding industries located in fewer than 20 counties.

### *A.2 Calculating location quotients of industries and clusters*

Our analysis starts with understanding the industrial structure and local strengths of all counties. First, we calculate LQs for NAICS six-digit industries for every county in the United States. LQ measures the local relative size of an industry:

$$LQ = \frac{\text{Proportion of industry employment in Local County}}{\text{Proportion of industry employment in Country}}. \quad (1)$$

LQs below 1 imply the industry is a smaller share of the county economy than the nation average. An LQ of 1 implies that a county has the same share of employment in that industry as the national average. LQs greater than 1 imply greater specialization in that industry. The QCEW already provides LQs, so for the relevant agriculture industries where the data is sourced from QCEW, we use the LQ as given.

### *A.3 Measuring relatedness densities of industries and clusters*

Relatedness industries are expected to frequently appear together. These places host all of the necessary capabilities, so co-location patterns reveal which industries require similar capabilities. Therefore, relatedness is measured by examining the probability that pairs of industries are located together in counties across the United States with LQs of at least 1. If two industries are likely to be located together and are rarely separate, it implies that they require very similar conditions to thrive, and are therefore highly related.

Thus, we measure relatedness between two industries by estimating the minimum conditional probability an industry has an LQ greater than 1, given that the other industry has an

LQ greater than 1.<sup>8</sup> The matrix  $M$  describes whether each place has an LQ greater than 1 in each industry, revealing it as a comparative advantage, and includes all regions and industries:

$$M = 1[|LQ_j \geq 1]. \quad (2)$$

From this matrix, we determine a co-occurrence matrix  $C$  that counts the number of places that have pairs of industries with LQs greater than 1:

$$C = M^T \cdot M \quad (3)$$

where the superscript  $T$  indicates that  $M$  is transposed. By definition, matrix  $C$  is symmetric and counts the number of times that each pair of industries appears together such that the diagonal (denoted  $C_{i,i}$ ) counts the number of occurrences of each individual industry. Dividing the number of co-occurrences for each pair of industries  $i$  and  $j$  by the maximum number of appearances for either industry estimates the minimum conditional probability that an industry has an LQ greater than 1, given that the other industry has an LQ greater than 1.

$$R_{i,j} = \frac{C_{i,j}}{\max(C_{i,i}, C_{j,j})} \quad (4)$$

This estimate captures the tendency for pairs of industries to cluster together in the same counties.

We then measure the intensity of relationships between each industry and the portfolios of industries in a specific county. Specifically, the relatedness density of industry  $i$  in county  $r$  is:

$$RD_{i,r} = \frac{\sum_j M_r \cdot R_{i,j}}{\sum_j R_{i,j}} \quad (5)$$

where  $M_r$  is the vector of industries in region  $r$  with LQs above 1. This measure of relatedness density captures the extent that related industries are present in a region, is widely shown to be predictive of an industry increasing its LQ above 1, and is the key indicator of feasible local industries in our study.

Finally, we also calculate the density of industry relatedness by cluster. We use the cluster definitions defined in Delgado et. al. (2016), which uses the principle of relatedness to group NAICS industries such that they are more related to industries within the cluster than outside the cluster. To incorporate agriculture and aquaculture from the QCEW data, we supplement the cluster definitions with a single agriculture cluster and add aquaculture to the Fishing and Fish Products cluster, renaming it Fishing and Aquaculture.

$$CD_{c,r} = \frac{\sum_{i \in c} RD_{i,r}}{\sum_{i \in c} R_{i,j}} \quad (6)$$

<sup>8</sup> Most calculations are performed using the EconGeo Package in R (Balland, 2017). Non-specialist readers may wish to skip the rest of this subsection describing the specific calculations.

This measure captures the relatedness density of all industries within a cluster that are hosted in a region relative to the relatedness of all pairs of industries. That is, a higher density of industry relatedness within clusters indicates the clusters that contain industries related to the region's existing industrial base.

#### *A.4 Data and methodological limitations*

The impact of data suppression on our calculations is ambiguous and may slightly underestimate relatedness at the industry-pair level, but we expect it has no effect on the overall level of relatedness density that we use to identify diversification opportunities. Specifically, data suppression affects both the numerator and denominator when calculating the conditional probability of an industry having an LQ greater than 1 and it has a slightly greater effect on the number of counties with both industries present than it has on the number of counties hosting one industry since the effect will be compounded by the requirement to host both. But relatedness density is a relative measure of the intensity of relatedness. So, if all industry pairs in the county are equally affected, there will be no general bias in the density measure. The issue diminishes further with a sufficient number of larger regions showing non-zero employment.

However, calculating density within a county only characterizes the intensity of relatedness to industries with an LQ greater than 1. In small counties where a significant number of industries have suppressed data it will therefore underestimate relatedness density. Specifically, density calculations will be affected by data suppression in smaller counties since suppressed industries would be excluded from each county's  $M_r$ , biasing density calculations downwards. However, the same  $M_r$  is used for all industries in the county, so all industries in the county can be expected to face the same downward bias. We identify diversification opportunities by examining the relative magnitudes of density within each county, so the downward bias in the smallest counties is generally not expected to bias conclusions. However, if a potential diversification opportunity is highly related to the industries with suppressed data, then it might not appear in the diversification opportunities even though it would have if the data were not suppressed. This risk is only a concern in smaller counties where data suppression is widespread, and practitioners should be cautious of results in very small counties. In counties with fewer industries with suppressed data, relatedness densities will be supported by the relatedness of non-suppressed industries and are still likely to appear in the results. Our case study is based on Hawai'i, which has relatively larger counties.

We also considered using more aggregate industry measures to reduce the effect of data suppression, such as three-digit NAICS codes. While this would have led to more accurate estimates of relatedness, the measure would reflect these larger categories and water down information about local capabilities. For regions, finding a specialized niche may be important, so disaggregated data offers more strategic alternatives. Previous studies show how different levels of aggregation reveal different information about those industries (Li & Neffke, 2024), rather than providing substitute measures. We examined various imputation methods to address missing data. Such methods are useful to minimize the effect of zeros in linear regression but would introduce additional biases into the relatedness estimates by assuming the presence of an industry, often when it is not present. So, the most accurate estimates of relatedness are still calculated with the raw data.

Since the CBP does not include agriculture, we use the QCEW to examine relationships with industries in agriculture. To protect confidentiality, the QCEW reports zeros in many

industries. This means some LQs will be shown as 0 in agricultural industries in many counties, when they are, in fact, not 0. The impact of zeros in the QCEW and data suppression in the CBP on our calculations is small, if any. Since relatedness is measured using a conditional probability, its calculation is conditional on counties reporting non-zero employment. Thus, the conditional probability estimated will still be accurate if there are a sufficient number of larger regions to accurately calculate the probabilities. However, it may bias the estimates of relatedness with agriculture, since statistical patterns for agriculture could differ in counties with larger populations. Nonetheless, the condition of non-zero employment already focuses the calculation on unaffected larger counties, so it would be detrimental to further limit calculations to a selection of counties or metropolitan areas. In any case, we are cautious about drawing conclusions for agriculture. We exclude self-employment and government data, which does not affect the calculation if the local proportion of self-employment in each industry is similar to the national proportion of self-employment in that industry.<sup>9</sup>

---

<sup>9</sup> A bias could occur if self-employment is proportional to the size of a region, but such a bias would be minimal as it only occurs around the threshold point where the LQ is equal to 1.