Economic Complexity of Australian Regions

Hamish Gamble

Sharif Rasel

2024-07-26

Abstract

## 1 Introduction

## 2 Literature Review

César A. Hidalgo and Hausmann (2009) introduced the concept of economic complexity as a means of quantifying and explaining differences in the economic development trajectory of different countries. Their method used bilateral trade data to identify the network structure of countries and the products they export and built on the concept of relatedness introduced in C. A. Hidalgo et al. (2007). Economic complexity has been shown to be a positive predictor of Gross Domestic Product (GDP), and GDP growth. Increasing economic complexity has also been shown to decrease unemployment and increase employment Adam et al. (2023), reduce green house gas emissions Romero and Gramkow (2021) and reduce income inequality Hartmann et al. (2017).

Relatedness has since been applied across industry (Neffke and Henning 2012), research areas (Guevara et al. 2016), occupation (Muneepeerakul et al. 2013) and technology (patents) (Kogler, Rigby, and Tucker 2013).

The relatedness approach has also been used to quantify economic complexity across cities, states, and regions, using employment dataChávez, Mosqueda, and Gómez-Zaldívar (2017), business counts(Gao and Zhou 2018), patent classifications (Balland and Boschma 2021), and interstate and international trade data (Reynolds et al. 2018).

Despite differences in data sources, the method for calculating economic complexity in the literature is relatively standard. The presence of an activity in a region is often identified using a location quotient method, such that an activity is said to be present in a region if:

Where is the measure of an activity in region - such as the level of employment in an occupation in a city, or the number of businesses classified in an industry in a province, or the value of exports of a product from a country. The location quotient method creates a binary matrix with rows and columns.

### 2.1 Regional economic complexity of small areas

The location quotient method can be unreliable due to the discontinuity at 1. This is especially relevant when economic complexity is calculated in regional areas where either or are small. In these cases, small changes, or measurement error in can significantly change the location quotient.

The choice of region size and activity classification is important. In a study of the economic complexity of US regions, Fritz and Manduca (2021) use metropolitan areas as the basis for calculations. Metropolitan areas in the United States are defined such that jobs within a given area are held by residents who live in that area.Metropolitan areas have a population of at least 50,000 people. The smallest MSA was estimated to have a 2023 population of 57,700 (about 0.015% of US population). They find a poor correlation between ECI calculated at higher level aggregated industry classifications indicating the importance of a high degree of disaggregation to provide as much information to the model as possible Fritz and Manduca (2021).

In New Zealand, Davies and Maré (2021) use weighted correlations of local employment shares. Regions range from a population of 1,434 to 573,150 with a mean population of 29,947 and median population of 6,952. Employment is measured as an industry-occupation pair.

* Differences in relationship between complexity and relatedness on indicators may be entirely context dependent.

## 3 Data & Methods

### 3.1 Data

* Calculate economic complexity indicators for Australian regions using employment data from the 2021 Census.
* Regions classified by Statistical Areas Level 3 (SA3)
* Economic activity classified by ANZSIC industry division and ANZSCO major group

Source: [Article Notebook](https://aiti-flinders.github.io/sirp-complexity/index.qmd.html)

* We exclude individuals who identify their place of work as a Migratory - Offshore - Shipping region or as No Fixed Address. Employment in these regions totals 497,913 or about 4% of the total sample.
* Following Davies and Maré (2021), employment is aggregated into industry-occupation pairs, allowing for differentiation between, for example, managers working in agriculture, forestry, and fishing, and managers working in retail trade.
* Dataset covers 340 regions and 152 industry-occupations. [Figure 1](#fig-employment-density) shows the presence of any level of employment within a region and industry-occupation. As can be seen, there is a high level of employment density across our data, with 88.5% of all combinations of region, industry, and occupation.

|  |
| --- |
| Figure 1: Presence of employment across regions and industry-occupations. |

Source: [Article Notebook](https://aiti-flinders.github.io/sirp-complexity/index.qmd.html)

### 3.2 Method

This section follows the method of Davies and Maré (2021) using correlations of employment shares rather than a location quotient method.

#### 3.2.1 Relatedness

Activities are related based on the weighted correlation between the local activity share of employment, weighted by each regions share of total employment.

* First calculate the weighted covariance
* Divide the weighted covariance by the city share-weighted standard deviations of the local activity shares to get the weighted correlation.
* Map the correlation to the interval such that:

City relatedness is calculated symmetrically such that:

#### 3.2.2 Complexity

Activity complexity is defined by the second eigenvector of the matrix and city complexity is defined by the second eigenvector of the matrix . The sign of activity complexity is set such that it is positively correlated with the weighted mean size of cities that contain activity , and the sign of city complexity is set such that it is positively correlated with the local share-weighted mean complexity of activities in city

Source: [Article Notebook](https://aiti-flinders.github.io/sirp-complexity/index.qmd.html)

## 4 Results

Source: [Article Notebook](https://aiti-flinders.github.io/sirp-complexity/index.qmd.html)

[Figure 2](#fig-gcc-complexity) shows the regional complexity of SA3 regions in Australian Greater Capital City Areas based on 2021 Census data. Complexity is highest in capital cities and surrounding regions.

|  |
| --- |
| Figure 2: Complexity of Australian Greater Capital City Areas |

Source: [Article Notebook](https://aiti-flinders.github.io/sirp-complexity/index.qmd.html)

### 4.1 Regression

Source: [Article Notebook](https://aiti-flinders.github.io/sirp-complexity/index.qmd.html)

### 4.2 Spatial Correlation

Source: [Article Notebook](https://aiti-flinders.github.io/sirp-complexity/index.qmd.html)

Source: [Article Notebook](https://aiti-flinders.github.io/sirp-complexity/index.qmd.html)

Based on [Figure 2](#fig-gcc-complexity), it looks like there are clusters of complexity, centred around capital cities.

|  |
| --- |
| Figure 3: Residuals from linear regression |

Source: [Article Notebook](https://aiti-flinders.github.io/sirp-complexity/index.qmd.html)

The residuals from the linear regression are shown in [Figure 3](#fig-residuals-map) which also shows that the distribution of the residuals appears non-random.

|  |
| --- |
| Figure 4: Moran Scatterplot for City Complexity in Australia |

Source: [Article Notebook](https://aiti-flinders.github.io/sirp-complexity/index.qmd.html)

The correlation between complexity and lagged complexity is shown in [Figure 4](#fig-moran-plot) which also shows a dependency. Finally, we observe a global Moran’s I of 0.498499 with a p.value of 0. As such, the data appear to be spatially autocorrelated, so a lagged AR or lagged error model should be estimated instead.

Source: [Article Notebook](https://aiti-flinders.github.io/sirp-complexity/index.qmd.html)

|  |  |  |  |
| --- | --- | --- | --- |
|  | OLS | Spatial AR | Spatial Error |
| (Intercept) | -17.224 \*\*\* | -13.830 \*\*\* | -16.581 \*\*\* |
|  | (3.417) | (3.351) | (3.561) |
| log\_total\_businesses\_no | -0.102 | 0.063 | -0.114 |
|  | (0.230) | (0.221) | (0.254) |
| log\_total\_business\_entries\_no | 0.028 | -0.060 | -0.037 |
|  | (0.409) | (0.389) | (0.395) |
| log\_total\_business\_exits\_no | 0.937 \* | 0.650 | 0.939 \* |
|  | (0.431) | (0.416) | (0.420) |
| turnover\_greater\_2m\_share | 0.491 | -0.083 | -0.220 |
|  | (1.748) | (1.668) | (1.721) |
| log\_total\_persons\_employed\_aged\_15\_years\_and\_over\_no | -1.691 \*\* | -1.382 \*\* | -1.793 \*\* |
|  | (0.517) | (0.500) | (0.565) |
| log\_estimated\_resident\_population\_persons\_no | 1.208 \* | 1.096 \* | 1.430 \* |
|  | (0.511) | (0.490) | (0.565) |
| business\_share | 4.280 \*\*\* | 4.457 \*\*\* | 4.371 \*\* |
|  | (1.206) | (1.149) | (1.415) |
| log\_median\_equivalised\_total\_household\_income\_weekly | 2.213 \*\*\* | 1.669 \*\*\* | 2.097 \*\*\* |
|  | (0.283) | (0.294) | (0.321) |
| rho |  | 0.273 \*\*\* |  |
|  |  | (0.060) |  |
| lambda |  |  | 0.360 \*\*\* |
|  |  |  | (0.066) |
| N | 330 | 330 | 330 |
| R2 | 0.528 | 0.560 | 0.574 |
| logLik | -346.835 | -338.084 | -335.202 |
| AIC | 713.669 | 698.168 | 692.403 |
| \*\*\* p < 0.001; \*\* p < 0.01; \* p < 0.05. | | | |

Source: [Article Notebook](https://aiti-flinders.github.io/sirp-complexity/index.qmd.html)

## 5 Hot spots

|  |
| --- |
| Figure 5: City Complexity hot spots (based on local Moran’s I p.values) |

Source: [Article Notebook](https://aiti-flinders.github.io/sirp-complexity/index.qmd.html)

## 6 Conclusion

## 7 Appendix

Regional economic complexity can be calculated using other data, including employment by industry and employment by occupation.

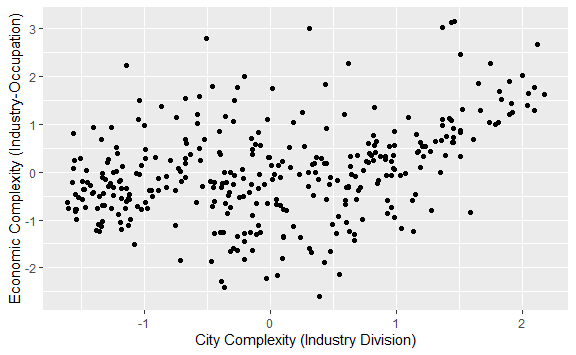
### 7.1 Other Employment Indicators

Source: [Article Notebook](https://aiti-flinders.github.io/sirp-complexity/index.qmd.html)

Employment density is much sparser when using 4-digit industry of employment data. Only 48 of the industry class-region combinations contain any level of employment.

|  |
| --- |
| Figure 6: Employment density, SA3 regions: (A) ANZSIC industry class, (B) ANZSCO occupation unit |

Source: [Article Notebook](https://aiti-flinders.github.io/sirp-complexity/index.qmd.html)



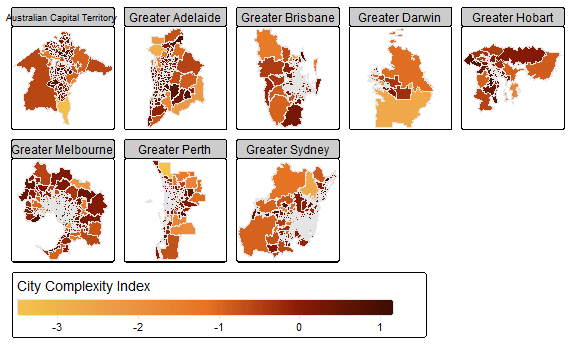
Source: [Article Notebook](https://aiti-flinders.github.io/sirp-complexity/index.qmd.html)

Source: [Article Notebook](https://aiti-flinders.github.io/sirp-complexity/index.qmd.html)

### 7.2 Smaller Areas

|  |
| --- |
| Figure 7: Employment density, Industry-occupation, SA2 regions |

Source: [Article Notebook](https://aiti-flinders.github.io/sirp-complexity/index.qmd.html)



Source: [Article Notebook](https://aiti-flinders.github.io/sirp-complexity/index.qmd.html)

## References

Adam, Antonis, Antonios Garas, Marina-Selini Katsaiti, and Athanasios Lapatinas. 2023. “Economic Complexity and Jobs: An Empirical Analysis.” *Economics of Innovation and New Technology* 32 (1): 25–52. <https://doi.org/10.1080/10438599.2020.1859751>.

Balland, Pierre-Alexandre, and Ron Boschma. 2021. “Mapping the Potentials of Regions in Europe to Contribute to New Knowledge Production in Industry 4.0 Technologies.” *Regional Studies* 55 (10-11): 1652–66. <https://doi.org/10.1080/00343404.2021.1900557>.

Chávez, Juan Carlos, Marco T. Mosqueda, and Manuel Gómez-Zaldívar. 2017. “Economic Complexity and Regional Growth Performance: Evidence from the Mexican Economy.” *Review of Regional Studies* 47 (2): 201–19. <https://doi.org/10.52324/001c.8023>.

Davies, Benjamin, and David C. Maré. 2021. “Relatedness, Complexity and Local Growth.” *Regional Studies* 55 (3): 479–94. <https://doi.org/10.1080/00343404.2020.1802418>.

Fritz, Benedikt S. L., and Robert A. Manduca. 2021. “The Economic Complexity of US Metropolitan Areas.” *Regional Studies* 55 (7): 1299–1310. <https://doi.org/10.1080/00343404.2021.1884215>.

Gao, Jian, and Tao Zhou. 2018. “Quantifying China’s Regional Economic Complexity.” *Physica A: Statistical Mechanics and Its Applications* 492: 1591–1603. https://doi.org/<https://doi.org/10.1016/j.physa.2017.11.084>.

Guevara, Miguel R., Dominik Hartmann, Manuel Aristarán, Marcelo Mendoza, and César A. Hidalgo. 2016. “The Research Space: Using Career Paths to Predict the Evolution of the Research Output of Individuals, Institutions, and Nations.” *Scientometrics* 109 (3): 1695–1709. <https://doi.org/10.1007/s11192-016-2125-9>.

Hartmann, Dominik, Miguel R. Guevara, Cristian Jara-Figueroa, Manuel Aristarán, and César A. Hidalgo. 2017. “Linking Economic Complexity, Institutions, and Income Inequality.” *World Development* 93 (May): 75–93. <https://doi.org/10.1016/j.worlddev.2016.12.020>.

Hidalgo, C. A., B. Klinger, A.-L. Barabaśi, and R. Hausmann. 2007. “The Product Space Conditions the Development of Nations.” *Science* 317 (5837): 482–87. <https://doi.org/10.1126/science.1144581>.

Hidalgo, César A., and Ricardo Hausmann. 2009. “The Building Blocks of Economic Complexity.” *Proceedings of the National Academy of Sciences* 106 (26): 10570–75. <https://doi.org/10.1073/pnas.0900943106>.

Kogler, Dieter F., David L. Rigby, and Isaac Tucker. 2013. “Mapping Knowledge Space and Technological Relatedness in US Cities.” *European Planning Studies* 21 (9): 1374–91. <https://doi.org/10.1080/09654313.2012.755832>.

Muneepeerakul, Rachata, José Lobo, Shade T. Shutters, Andrés Goméz-Liévano, and Murad R. Qubbaj. 2013. “Urban Economies and Occupation Space: Can They Get “There” from “Here”?” Edited by César A. Hidalgo. *PLoS ONE* 8 (9): e73676. <https://doi.org/10.1371/journal.pone.0073676>.

Neffke, Frank, and Martin Henning. 2012. “Skill Relatedness and Firm Diversification.” *Strategic Management Journal* 34 (3): 297–316. <https://doi.org/10.1002/smj.2014>.

Reynolds, Christian, Manju Agrawal, Ivan Lee, Chen Zhan, Jiuyong Li, Phillip Taylor, Tim Mares, Julian Morison, Nicholas Angelakis, and Göran Roos. 2018. “A Sub-National Economic Complexity Analysis of Australia’s States and Territories.” *Regional Studies* 52 (5): 715–26. <https://doi.org/10.1080/00343404.2017.1283012>.

Romero, João P., and Camila Gramkow. 2021. “Economic Complexity and Greenhouse Gas Emissions.” *World Development* 139 (March): 105317. <https://doi.org/10.1016/j.worlddev.2020.105317>.