1 Question 1

> **It is commonly assumed that agglomeration economies generate benefit to: i) the firms within a cluster and ii) the communities that 'house' the agglomeration. In both cases, provide detail on the nature of these advantages.**

Lean heavily on Capello (2007) taxonomy of advantages.

Production Cost Co-location in a cluster offers many advantages to both the firm and the community that housed the cluster, such as transportation cost

Firms can lower costs by minimizing transportation costs between steps of production. The classic extreme case is the Ford Motor Company's River Rouge complex that took in raw materials from vertically integrated subsidiaries such as iron ore, coal, wood and rubber at one end, and delivered finished cars at the other end. However, due to management costs, the set-up was not as efficient as originally envisioned by the theoreticians of scientific management. A more modern example is the current Western just in time system (JIT/lean manufacturing), that emphasizes close integration between parts suppliers and assembly/sub-assembly plants since there is a cost associated with each manipulation of inventory and firms attempt to keep a minimum of inventory at hand at each step of manufacturing (Hallihan, Sackett, and Williams 1997). In addition, shorter distances involved between parts suppliers and assembly locations results in more trips per unit of time, and thus firms save directly on transportation costs. (Heuvel, Langen, Donselaar, and Fransoo 2014)

Furthermore, in areas with access to temporary skilled-workers, companies are able to ramp up and slow down production to match the business environment without having any long-term association/relationship to the labour. As seen in the current "gig" economy, employers are able to externalize the cost of down-time onto labour, thereby reducing cost. Furthermore, this high level of mobility facilitates the transfer of tacit knowledge (best practises) between firms.

A major community advantage of being located in a cluster, is the abilit to change employers in the same field of work, without having to move house and home (Hickling Arthurs Low 2009). This helps preserve communal bonds, while lowering the transaction costs of changing employers.

2 Question 2

Discuss possible conventional GIS modeling of some of your data variables (pp. 14-15 in your Proposal). This may include Overlay Combinations, Rules in Distance Relationships, Spatial Analysis using Continuous Fields and other functions, normally as part of GIS procedures (scripts). Allude to both (quantitative) Analysis and (qualitative) Visualization results. Specify the spatial scope for these models, e.g., within cities, regions (GTA, SW Ontario), or at coarse scales (Canada, USA, North America).

Descriptive: Quadrat Analysis would use the geocoded points of the institutional investor location in the GIS environment. and aggregate the number of points in each pixel of the quadrat lattice. However, this raster will be quite noisy. Therefore, to simplify the analysis of aggregate weight of institutional investors, a median filter will be used to smooth the data. A median filter is more likely to be of greater use than a Fast-Fourier Transform since the data is more likely than not to be non-parametric - a key assumption behind Fourier Transform - as well having the property of preserving edges. This new surface can be used to identify local concentrations of investors if using a gird on the order of a few hundred meters, as well as continental scales if using girds that covers hundreds of kilometres per side.

Heat maps are a useful for discerning point patterns at various scales from continental (Canada and the United States) to regional (Central Canada[1], Northeastern United States[2])neighbourhoods (Downtown Toronto, Manhattan). There are many ways to calculate the heatmap, from simple counts to weighing according to various variables, such as money under management. While not often considered an overlay operation, heat maps are overlay operations behind the scenes, since the algorithm calculates a kernel (2-d Gaussian probability distribution around the point) and sum all of the layers of the operation into a final heat map where a density of points translates into higher values due to the cumulative effect.

[1] Ontario and Quebec

[2] This region consists of : Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania Rhode Island, and Vermont (United States Department of Commerce and Statistics Administration 2010)

Spatial analysis routines

Since the coverage area remains the same for the entire study period, it is possible to examine changes in test results over time without having to worry about the neighbourhood effect since physical space remains constant over time. This would be a great opportunity to look at large North American cities such as New York, Boston, Chicago, Toronto, Philadelphia and San Francisco and test a few of the conclusions set out in Bodenman (1998) and Bodenman (2000), such as the exodus of algorithmic traders/index funds trader to the city periphery. This can be accomplished by aggregating the total funds allocated to different indexes (Dow-Jones 100, S&P 1000, FTSE 1000 and TSX 1000) for each year 2010 census block. In order to keep spatial boundaries constant throughout the study period (2000 to 2015) the Year 2010 census block shapefile will be used for spatial aggregating rather than using a combination of the year 2000 and 2010 shapefile. These 64 populated shape files will be transformed into a movie in order to facilitate the tracking of movement over time. A second exploratory method for analyzing the spatio-time series data would be to apply Tukey's median polish to a matrix comprised of polygon IDs on one axis and time intervals on the other (Tukey 1977; Martinez and Melo 2015). By analyzing the residuals table of Tukey's median polish algorithm, in conjunction to the movie created in the previous step will highlight areas interest for more informed analysis rather than simple exploratory data analysis.

Nearest Neighbour Analysis Ripley's K

Spatial Shift-Share Analysis (Gian Pietro Zaccomer and Luca Grassetti 2008)

3 Question 3

> **Discuss the concept of spatial autocorrelation. Explain and critically evaluate that concept in the context of the spatial patterns of economic activities.**

Spatial autocorrelation is at it's heart the application of Tobler's first law of geography: "everything is related to everything else, but near things are more related than distant things."(Tobler 1970, page 236). In other words, this concept test using probability on whether the characteristics of location A are similar (or dissimilar) to location B via distance lags. Significant results implies that spatial processes are at work. Furthermore, this underlying spatial pattern renders many classical statistics unreliable, since it violates the assumption of independence between observations(Fischer and Getis 2009).

There are four widely used measures of spatial autocorrelation: Joint count analysis, Moran's I, Geary's C and Getis-Ord g statistic (also known in the literature as Hotspot Analysis or High/Low Clustering) as well as Local Moran's I.

The simplest test for spatial autocorrelation is joint count analysis (Lee and Wong 2000). This technique measures the strength of association of a phenomenon with regards to presence or absence of a binary attribute in adjacent polygons (ie sum of 1-1, 0-1, 0-0 interactions) vs an expected count that assumes the distribution of polygon counts is randomly distributed. As with many spatial techniques, this measure is sensitive to the border effect since unnecessarily large survey areas or unessessarily small polygons (using dissemination blocks for Canada-wide data location of Tim Hortons franchise locations) can artificially increase the effect of clustering. The reverse is also true, with too large a spatial aggregation reducing the power of the test to irrelevance (ie. using presence in forward sortation areas polygons).

While this method is simple (conceptually and mathematically), it's simplicity can make it a coarse instrument in some respects. For instance, the binary attribute (presence or absence), does not measure the strength of association, therefore a polygon with 1 instance is weighted the same as a polygon with 10 instances. In order to calculate this level of interaction, a more sophisticated measure will be needed, such as Moran's I statistic.

Moran's I, is more sophisticated than joint-count analysis, by allowing the use of points and polygons as well as the ability to use any continuous variable rather than binary data. The output of Moran's I measures the degree of correlation between the variable χ and the spatial lag of χ by averaging the value of χ in neighbouring regions as determined by the spatial weights matrix.

Furthermore, the spatial weights table is more customizable to the problem at hand than in Joint-Count Analysis. Most of the decision process in choosing the distance method for populating the spatial weights table will be influence by the literature. From the most naive (least subject knowledge/fewest assumptions), the spatial weights table can describe a basic agency matrix (1 or 0) using the rook's or queen's case when dealing with polygons, or in the case of point data, the nearness of the point/polygon to it's neighbours ($wij = 1/dij$ where the matrix is calculated as 1 over the distance of between points) for n-number of neighbours or buffer of n-distance from each point/centroid of polygon. Furthermore, the adjacency matrix can also be row standardized when there is an unequal number of neighbours, or have a distance decay function such as exponential decay.

With regards to spatial pattern of economic activity, Moran's I can be used to test whether certain industries cluster or reppel each other. For example, Green, O'Hagan, and Lefebvre (2015) shows that institutional investors located on the island of Manhattan show clustering behaviour by co-locating more than what would be expected by random chance. However, a more sophisticated treatment of Moran's I by mesuring the association between the weighted return on capital metric rather than simple presence, and found that metric was randomly distributed. This makes conceptual sense, since it is unlikely that a potentially footloose industry (very few sunk costs with regards to location when compared to heavy industry) investors could profit by taking advantage of spatial location alone on an ongoing period, especially when the dominant investing paradigm indicates that it is impossible to beat the market on a long-term basis (Malkiel 2005).

More broadly, while Moran's I can be used to test for the presence of spatial autocorrelation of an industry within a certain geographical area, it's can't pinpoint local clusters. This is where the localized variant of Moran's I, also known as Local Indicators of Spatial Association (LISA) should be used (Anselin 1995). This variant of Moran's I simultaneously calculates the degree of spatial autocorrelation between one point/polygon to all neighbouring points/polygon in order to identify each point/polygon as belonging to one of 4 states (High/High, Low/Low (Similarity to neighbours), High/Low and Low/High(dissimilar to neighbours/possible outliers) as well as indicating the absence of significant relationships (Anselin 2003). With regards to using LISA with economic data, one could do worse than using this technique to find clusters of economic agglomeration. For example, if analyzing the spatial autocorrelation of firm headquarters, cluster such as Toronto's Financial District would likely show up as a High-High district, whereas nearby Regent's Park would likely be a Low-High district (near Downtown, but relatively few/non-existent amount of headquarters compared to it's environs) .

References

Anselin, L. (1995). Local Indicators of Spatial Association – LISA. *Geographical analysis 27*(2), 93–115.

Anselin, L. (2003). *GeoDa^{TM}0.9 User's Guide*. Spatial Analysis Laboratory Department of Agricultural and Consumer Economics University of Illinois, Urbana-Champaign. `https://geodacenter.org/downloads/pdfs/geoda093.pdf`.

Bodenman, J. E. (1998). The Suburbanization of the Institutional Investment Advisory Industry: Metropolitan Philadelphia, 1983–1993. *The Professional Geographer 50*(1), 112–126.

Bodenman, J. E. (2000). Firm Characteristics and Location: The Case of the Institutional Investment Advisory Industry in the United States, 1983 – 1996. *Papers in Regional Science 79*(1), 33–56.

Capello, R. (2007). *Regional Economics*. Routledge.

Fischer, M. M. and A. Getis (2009). *Handbook of applied spatial analysis: software tools, methods and applications*. Springer Science & Business Media.

Gian Pietro Zaccomer and Luca Grassetti (2008). Using R for Spatial Shift-Share Analysis. In *UseR!2008, Dortmund, Germany*. `https://www.r-project.org/conferences/useR-2008/slides/Zaccomer+Grassetti.pdf`.

Green, M. B., S. O'Hagan, and M. Lefebvre (2015). The economic geography of institutional investment in the united states, 2010. *Urban Geography 36*(6), 798–822.

Hallihan, A., P. Sackett, and G. Williams (1997). JIT manufacturing: The evolution to animplementation model founded in current practice. *35*(4), 901–920.

Heuvel, F. P., P. W. Langen, K. H. Donselaar, and J. C. Fransoo (2014). Proximity matters: synergies through co-location of logistics establishments. *Journal of Logistics 17*(5), 377–395.

Hickling Arthurs Low (2009). Canada's entertainment software industry: The opportunities and challenges of a growing industry. `http://theesa.ca/wp-content/uploads/2015/08/Economic-Impact-report-2009.pdf`.

Lee, J. and D. W.-S. Wong (2000). *GIS and Statistical Analysis with ArcView.* John Wiley. 208 pages.

Malkiel, B. G. (2005). Reflections on the efficient market hypothesis: 30 years later. *40*(1), 1–9.

Martinez, W. and C. Melo (2015). *STMedianPolish: Spatio-Temporal Median Polish*. R package version 0.1.

Tobler, W. R. (1970). A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography 46*, pp. 234–240.

Tukey, J. W. (1977). *Exploratory Data Analysis* (First Edition ed.). Addison-Wesley Publishing Company. 688 pages.

United States Department of Commerce, E. and G. D. Statistics Administration, United States Census Bureau (2010). Census Regions and Divisions of the United States. Available online at `https://www2.census.gov/geo/pdfs/maps-data/maps/reference/us_regdiv.pdf`.