London Underground Network Analysis

I. Topological network

I.1. Centrality Measures

We use the degree centrality, closeness centrality, and betweenness centrality methods to measure node importance. Degree centrality represents how well a node is connected to other nodes. Degree centrality $C_D(u)$ is given by equation below, where n is the total number of nodes of the network. (NetworkX, no date)

$$C_D(u) = \frac{\text{Number of edges of u}}{n - 1}$$

In the underground network, the ranking of it depends on the number of connections the station has with other stations. The more underground lines pass through a station, the greater the $C_D(u)$, and the more important that station is.

Table 1: Top 10 most important stations by degree centrality

	Degree Centrality
Stratford	0.0225
Bank and Monument	0.0200
Baker Street	0.0175
King's Cross St. Pancras	0.0175
Oxford Circus	0.0150
Canning Town	0.0150
Green Park	0.0150
Liverpool Street	0.0150
Waterloo	0.0150
Earl's Court	0.0150

Closeness centrality determines the importance of a node by measuring the average distance of each node from the other nodes. (NetworkX, no date) The smaller average distance, the more important it is.

$$C_B(u) = \frac{n-1}{\sum_{v=1}^{v=n-1} d(v,u)}$$

In underground networks, the value is the reciprocal of the average number of stops required to get from one station to all other stations, and a larger Closeness centrality means that the station is more central and closer to other stations in the network. For example, Green Park has the highest closeness centrality of 0.114778, implying it requires an average of 8.7125 stops to get to all other underground stations.

Table 2: Top 10 most important stations by closeness centrality

	Closeness Centrality
Green Park	0.114778
Bank and Monument	0.113572
King's Cross St. Pancras	0.113443
Westminster	0.112549
Waterloo	0.112265
Oxford Circus	0.111204
Bond Street	0.110988
Angel	0.110742
Farringdon	0.110742
Moorgate	0.110314

Betweenness centrality determines the importance of a node by measuring how often the node is on the shortest path between other stations, shown in following equation. (NetworkX, no date) Where l denote the number of best paths.

$$p_{st}^{u} = \begin{cases} 1 \text{ if shortest path from s to t path through u} \\ 0 \text{ otherwise} \end{cases}$$

$$c_B(u) = \sum_{s,t \in G} \sum_{l} \left(\frac{p_{st}^{u}}{p_{st}}\right)_l$$

Betweenness centrality implies the proportion of the shortest paths that a station takes between any two stations, a higher betweenness centrality means that this station is likely the 'intermediary' between the stations. The efficient connection between two stations in the network depends to a large extent on this station functioning properly.

Table 3: Top 10 most important stations by betweenness centrality

	Betweenness Centrality
Stratford	0.297846
Bank and Monument	0.290489
Liverpool Street	0.270807
King's Cross St. Pancras	0.255307
Waterloo	0.243921
Green Park	0.215835
Euston	0.208324
Westminster	0.203335
Baker Street	0.191568
Finchley Road	0.165085

I.2. Impact measures

We will use Transitivity and Global Efficiency to measure the impact of removing nodes on the network. Transitivity, also known as global clustering coefficient, focuses on the triads in the network to measure the connectivity. It is a value between 0 and 1 and with the better-connected networks converging to 1. (NetworkX, no date)

$$\mathrm{T} = \frac{\text{Number of triangles } \times 3}{\text{Number of connected triplets}}$$

Transitivity can be used to assess the connectivity tightness, or aggregation, of most networks. For example, this metric can measure the density of connections between friends of friends across a social network, thus reflecting whether the network is a 'circle of friends'.

We measure the efficiency of London Underground network using Global Efficiency, which represents the inverse of the average shortest distance between all nodes in the network. (NetworkX, no date)

$$\mathrm{GE} = \frac{\mathrm{n} \times (\mathrm{n} - 1)}{\sum_{\mathrm{s,t} \in \mathrm{G}} \mathrm{l}_{\mathrm{st}}}$$

A shorter average distance means that the network has higher efficiency, and any two stations are more accessible. When stations at the edge of the network are removed, Global Efficiency may increase, meaning that the stations are closer together, but ignoring the accessibility of the edge areas. When important stations are removed from the network, Global Efficiency should be monotonically decreasing, with the absence of hub stations reducing the efficiency of the network.

Global Efficiency can assess the efficiency of information transfer and resilience of other networks. In the case of a company's employee network, more connectivity between employees will increase the efficiency of information transfer and thus improve communication. While companies with high Global Efficiency are less likely to experience downtime when key managers leave the company.

I.3. Node removal

Considering the two impact measures of Transitivity and Global Efficiency, and the two strategies of non-sequential and sequential, we believe that betweenness centrality can better reflect the importance of underground stations in the network.

As discussed earlier, betweenness centrality calculates the proportion of underground stations that appear in the shortest path between all stations on the network, using this metric we can more accurately estimate the impact of closing the station on the system. When the betweenness centrality of a station is high, its accessibility largely controls the flow of traffic, and if it is closed the system will lose its important 'intermediary'. If it is the only 'bridge', passengers will not be able to reach their destination via the underground and must choose a longer ride to bypass the station. In contrast to degree centrality, which focuses on the number of connections, there may be stations where the intersection of two lines is more important than the intersection of three highly alternative lines because it is the only link between the two or an important hub which the degree centrality does not reflect well.

As reflected in Figure 2.2 global efficiency data, the removal of the first five most important stations has a similar impact on the network for both degree centrality and betweenness centrality, but further it is clear that degree centrality does not reflect more important stations well. The number of stations is not a good indicator of the importance of the underground stations, whereas betweenness and closeness centrality can be relatively more accurate. The closeness centrality has a serious drawback in that when the network is not fully connected as important stations are removed, the average distance from each station to all stations in the same sub-network does not accurately reflect the importance of that station.

Figure 1.1: Transitivity of Centrality Non-sequential Removal

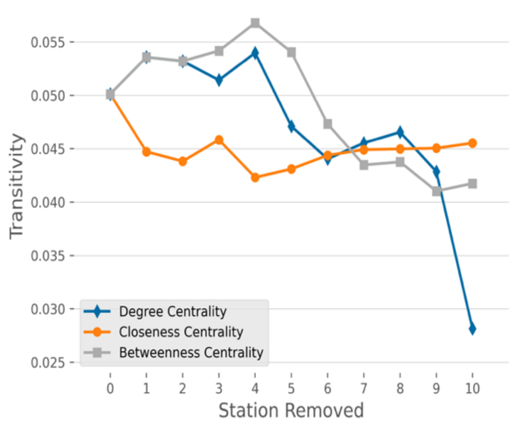


Figure 1.2: Global Efficiency of Centrality Non-sequential Removal

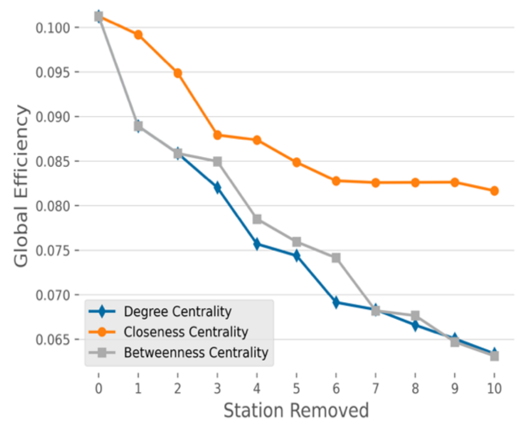


Figure 2.1: Transitivity of Centrality Sequential Removal

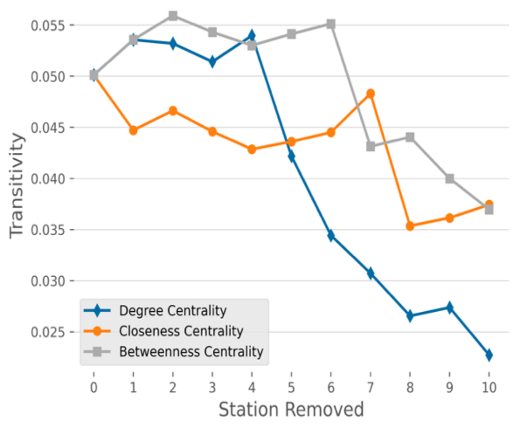
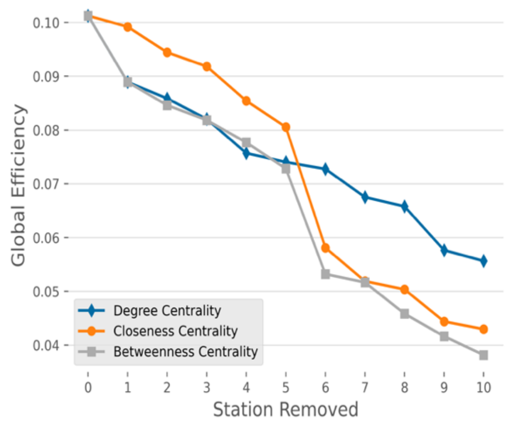


Figure 2.2: Global Efficiency of Centrality Sequential Removal



Sequential removal is more advantageous than non-sequential removal in testing the resilience of the network. For the non-sequential removal strategy, we start by calculating the most important underground stations in the case of a complete network, and the importance of stations in the network changes as stations are removed. The sequential removal strategy recalculate the current most important station to ensure that we are damaging the current network most efficiently.

Using Table 4 as an example, for transitivity, the more important stations for the connectivity of the network are removed in sequential strategy. For global efficiency, sequential removal reduces it to lower and it requires a longer distance to reach other stations. Therefore, sequential removal can be more accurate in testing the resilience of the underground network by dynamically determining critical stations.

Table 4: Impact measures of removing 10 stations for the betweenness centrality

	Non-sequential Removal	**Sequential Removal**
Final transitivity after 10 removals	0.041739	0.036972
Final global efficiency after 10 removals	0.063139	0.038164

Global efficiency can be better used as a global variable than transitivity to determine the loss of network connectivity when stations are removed. Transitivity is limited to the interconnection of stations direct connections, which can be used as an indicator of how well connected the network is, but it is uncommon to see three interconnected stations. More detailed, transitivity may rise or fall when a node is removed from the network. Transitivity may increase when the removal of a node causes triads to fall while the number of triangles remains the same, or when triads fall faster than triangles, but in most cases, transitivity follows a trend of decreasing as nodes are removed. In contrast, global efficiency, by measuring the average distance between all nodes in the network, provides a more accurate picture of the impact when some important nodes are removed.

Secondly, global efficiency is monotonically decreasing regardless of whether the non-sequential or sequential approach is used, which let us quantify the impact of stations removal more directly. Finally, global efficiency is more practical than transitivity, as the average distance between sites is more intuitive than triads in terms of the loss to the network due to the removal of nodes.

II. Flows: weighted network

II.1.

Degree centrality focuses on the number of connections between stations, which is determined when the infrastructure is established, thus not affected by passenger flows. In weighted network closeness centrality, we abstractly interpret the flow as distance, where a high flow between two stations is considered close using the inverse of flow. A station with high traffic will be considered closer to the center of the network with higher betweenness centrality. Similarly, for betweenness centrality, we use the flow to calculate the weighted shortest path. A station with the high flow is closer to the surrounding stations and more likely to be the shortest path, thus increasing the betweenness centrality.

Table 5: Top 10 most important stations by weighted closeness centrality

	Weighted Closeness Centrality
Green Park	95.978788
Westminster	95.969740
Waterloo	95.962234
Bank and Monument	95.959071

Oxford Circus	95.943574
Victoria	95.932390
Bond Street	95.928838
Liverpool Street	95.924381
Warren Street	95.900359
Moorgate	95.891873

Table 6: Top 10 most important stations by weighted betweenness centrality

	Weighted Betweenness Centrality
Green Park	0.568784
Bank and Monument	0.505050
Waterloo	0.414098
Westminster	0.378308
Liverpool Street	0.336316
Stratford	0.331378
Euston	0.284236
Oxford Circus	0.257143
Bond Street	0.256053
Baker Street	0.250602

II.2.

Transitivity does not consider the weight of edges, so for weighted networks we use the average clustering coefficient which reflects how tightly connected the nodes are. The weighted clustering coefficient of all nodes in the network is calculated, where the two edges passing through node u, uv, and uw, and vw form a triangular structure, and w is the weight of each edge, the flow. (NetworkX, no date)

$$c_u = \frac{1}{\deg(u)(\deg(u)-1))} \sum_{v,w\in G, \exists uv,uw,vw} (\hat{w}_{uv}\hat{w}_{uw}\hat{w}_{vw})^{1/3}$$

$$\hat{w}_{st} = \frac{w_{st}}{\max(w)}$$

$$C_G = \frac{1}{n} \sum_{u \in G} c_u$$

We use the average shortest path length, weighted by the inverse of the flow, which calculate the weighted average distance between nodes considering passenger traffic. (NetworkX, no date)

$$\bar{L} = \sum_{\substack{s,t \in G \\ s \neq t}} \frac{d(s,t) \times \frac{1}{flow_{s,t}}}{n(n-1)}$$

II.3.

According to I.1. we used betweenness centrality to select the three most important stations and used sequential removal. Table 7 below shows the outputs.

Table 7: Impact of London underground station removal

Remove Station	Remain station	Transitivity	Average Clustering	Global Efficiency	Average Shortest Path Length
	401	0.0501	0.1633	0.1013	2.0733
Green Park	400	0.0447	0.1400	0.0992	2.0807
Bank and Monument	399	0.0438	0.1170	0.0949	2.0907
King's Cross St. Pancras	398	0.0458	0.1178	0.0879	2.1118

We found that as the station was removed both transitivity and average clustering coefficient, two variables that focus on the triangle structure, gradually decreased, with a small fluctuation when King's Cross St. Pancras was removed. For the weighted average clustering coefficient, the slight increase means that King's Cross has a lower clustering coefficient than the average. This again supports the point made earlier that it may not be a good choice to use the triangle structure to determine the connectivity.

The depth of the red fill of the cells represents the corresponding index fluctuations due to the removal of the station, with the larger the fluctuations the darker the color. We find that the removal of Bank and Monument results in the largest fluctuations in the average clustering coefficient, so the triangle structure suggests that the closure of Bank and Monument would have the greatest impact on commuters, considering flow.

While global efficiency decreases and the average shortest path length increases monotonically, the weighted average distance between stations gradually increases as important stations are removed. When a high-traffic key station is removed from the network, the shortest path between other stations has one less path with a shorter 'distance'. From the coloring of the cells, we can see that the removal of King's Cross has a large fluctuation in the global efficiency and average shortest path length of the network and is therefore considered to have the greatest impact on passengers.

III. Models and calibration

III.1.

The spatial interaction models are derived from the basic gravitational model to measure the flow of traffic, population, economy, etc. (Wilson, 1971) The basic models are as follows:

$$T_{ij} = K\frac{O_i D_j}{d_{ij}^{\beta}}$$

Where K is the scale constant, controlling the overall flow. O_i, D_j represents the origin and destination attraction, which can be population, jobs, or GDP. d_{ij} represents the hindering factors between origin i and destination j, which can be the distance between, transportation cost, etc., and controlled by the coefficient β. And T_{ij} representing the flow between origin i and destination j.

When the values of attractiveness represented by O_i, D_j cannot be compared directly, we add the exponential to O_i, D_j. and the hindering factor $c_{ij}^{-\beta}$ can also be replaced by the negative exponential $\exp(-\beta c_{ij})$. We obtain the new spatial interaction model as follows:

$$T_{ij} = KO_i^{\alpha}D_j^{\gamma}\exp(-\beta c_{ij})$$

Also, we have the constrain that the sum of the model prediction and the actual flow is equivalent:

$$T = \sum_i \sum_j T_{ij}$$

Taking T_{ij} into the formula gives:

$$K = \frac{T}{\sum_i \sum_j O_i^{\alpha}D_j^{\gamma}\exp(-\beta c_{ij})}$$

Thus, in the model that restricts only the sum of flows, the flow is:

$$T_{ij} = T\frac{O_i^{\alpha}D_j^{\gamma}\exp(-\beta c_{ij})}{\sum_i \sum_j O_i^{\alpha}D_j^{\gamma}\exp(-\beta c_{ij})}$$

It is called unconstrained model since the out-bound flow at the origin and the in-bound flow at the destination are not constrained.

Sometimes we only know the out-bound flow at the origin, so we need to constrain the flow at the origin, which is the Origin-Constrained Model. K is replaced by the coefficient A_i, which is constrained by the origin flow.

$$A_i = \frac{1}{\sum_j D_j\exp(-\beta c_{ij})}$$

$$T_{ij} = A_iO_iD_j\exp(-\beta c_{ij})$$

$$\sum_j T_{ij} = O_i$$

The Destination-Constrained Model can be built in a similar way when we only know the in-bound flow of the destination.

$$T_{ij} = O_iB_jD_j\exp(-\beta c_{ij})$$

$$\sum_{i} T_{ij} = D_j$$

When both the departure and destination flows are known, we can add both constrain to the model, which is called the Doubly Constrained Model.

$$T_{ij} = A_i O_i B_j D_j \exp(-\beta c_{ij})$$

$$\sum_{j} T_{ij} = O_i$$

$$\sum_{i} T_{ij} = D_j$$

III.2.

We use the population of the departure stations, the number of jobs at the destination stations, and the distance between stations provided in the London Underground flow data to build the four spatial interaction models.

To obtain a linear relationship between the independent and dependent variables, we take the logarithm of the population, jobs, and distances. Since the flow between stations follows Poisson distribution, all flows are non-negative and have right-skewed, we fit the Poisson regression. We predict the flow using the parameter λ, which represents the expectation of the Poisson distribution. using the Unconstrained Model as an example:

$$\lambda_{ij} = \exp\left(K + \alpha \cdot \ln O_i + \gamma \cdot \ln D_j - \beta \cdot \ln c_{ij}\right)$$

We calculate the R-squared and root mean square error between the predicted and actual values for models. R-squared represents the degree of explanation of predicted value to actual value in variation, while RMSE represents the error. In Table 8, we find that the Doubly Constrained Model have better performance.

Table 8: R-Squared and RMSE of spatial interaction model

Model	R-Squared	RMSE
Unconstrained	0.3212	108.33

Origin-Constrained	0.3883	102.89
Destination-Constrained	0.3499	106.01
Doubly Constrained	0.4077	101.34

Since the population at the origin and the number of jobs at the destination have the same magnitude, the exponent set to 1. By comparing the results of using $\beta \cdot \ln c_{ij}$ and $\beta \cdot c_{ij}$ as the distance suppression function, we find that $\beta \cdot c_{ij}$ performs better, thus our spatial interaction model is:

$$T_{ij} = A_i O_i B_j D_j c_{ij}^{-\beta}$$

$$\sum_j T_{ij} = O_i$$

$$\sum_i T_{ij} = D_j$$

$$A_i = \frac{1}{\sum_j B_j D_j c_{ij}^{-\beta}}$$

$$B_j = \frac{1}{\sum_i A_i O_i c_{ij}^{-\beta}}$$

$$\lambda_{ij} = \exp(A_i + B_j - \beta \cdot c_{ij})$$

Finally, since A_i and B_j in the model depend on each other, we tried to improve the model by iteratively computing them using entropy approach (Senior, 1979), but the predictive performance does not improve. Thus, we have:

$$T_{ij} = \exp(A_i + B_j - 0.0001543 \cdot c_{ij})$$

IV. Scenarios

IV.1.

To simulate the impact of reductions in jobs at Canary Wharf on London Underground flows, an Origin-Constrained Model was developed, where the control coefficient A_i ensures that the departure flow from each station remains constant, D_j represents the number of jobs at the destination station j, c_{ij} represents the distance between stations i and j, and β was set to 0.00015.

$$\lambda_{ij} = \exp\left(A_i + \gamma \cdot \ln D_j - \beta \cdot c_{ij}\right)$$

$$A_i = \frac{1}{\sum_j D_j \exp(-\beta c_{ij})}$$

We build the baseline using the original number of jobs at Canary Wharf and the predictive model using the data after the halving of jobs. Comparing the before and after London Underground flows analyses the impact. Table 9 shows the top 10 Tube stations with the largest reduction and increase in destination flows before and after the jobs reduction.

Table 9: London Underground Stations Inbound Flow Change

Station	Reduction
Canary Wharf	-18,060
Romford	-15
Ilford	-14
Chadwell Heath	-13
Goodmayes	-12
Harold Wood	-11
London Fields	-11
Seven Kings	-11
Finchley Road & Frognal	-10
Shenfield	-10

Station	Increase
Bank and Monument	2,263
Stratford	1,939
Liverpool Street	968
Canning Town	724
Oxford Circus	582
King's Cross St. Pancras	429
London Bridge	416
Victoria	395
Farringdon	343
Highbury & Islington	343

We found that the main reduction is from Canary Wharf, with fewer passengers traveling there as jobs were lost, and an increase in flows to Bank and Monument, Stratford station. With total flows and departure flows remaining, this implies a shift in passenger destinations from Canary Wharf to other stations. Figure 3 shows the reduction in flows destined for Canary Wharf from other stations, and we can see that most of the reduction is concentrated around Canary Wharf, in the eastern part of the city, and along the lines connecting to Canary Wharf. Table 10 shows detailed data.

Figure 3: Outbound Flow Reduction in London Underground Stations

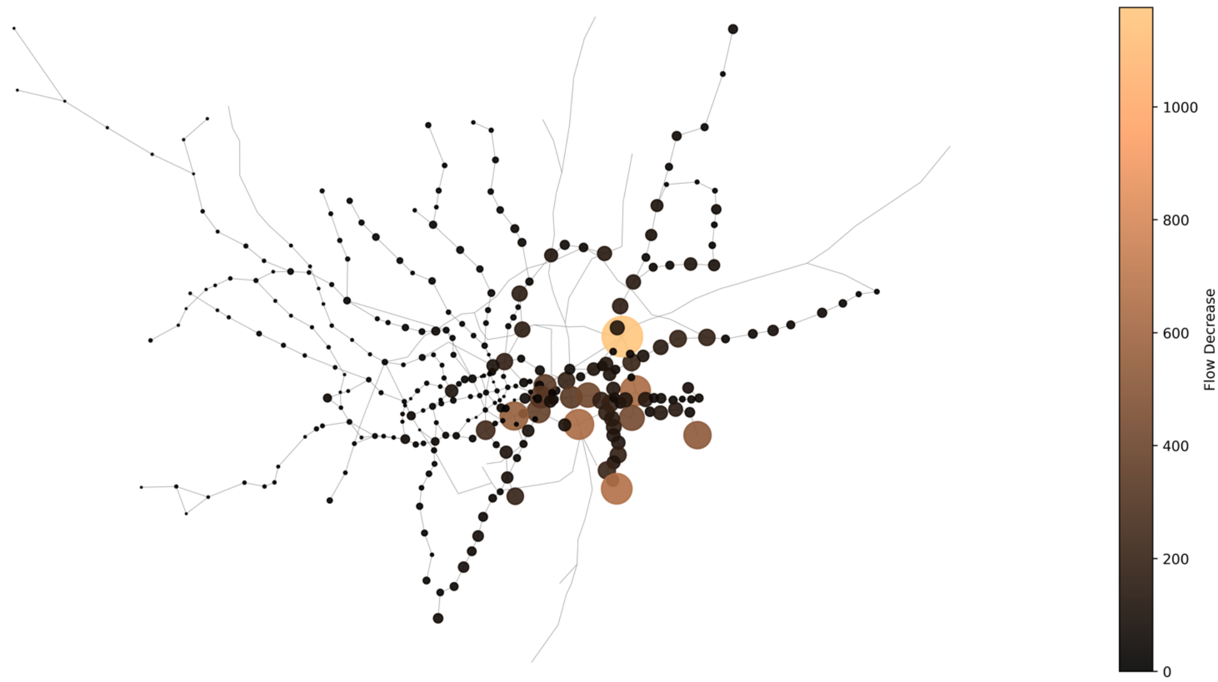


Table 10: London Underground Stations Outbound Flow to Canary Wharf Reduction

Station	Prior	Post	Reduction	Reduction Percentage
Overall	47681	29621	-18060	-38%
Stratford	3002	1826	-1176	-39%
Lewisham	1981	1314	-667	-34%
Canning Town	1655	1017	-638	-39%
Canada Water	1623	989	-634	-39%
Waterloo	1400	840	-560	-40%
Woolwich Arsenal	1481	955	-526	-36%
North Greenwich	1140	707	-433	-38%
Limehouse	1217	798	-419	-34%
London Bridge	923	555	-368	-40%

IV.2.

We build the doubly constrained model to simulate the effect of transport costs increase.

$$\lambda_{ij} = \exp\left(A_i + B_j - \beta \cdot c_{ij}\right)$$

The cost function coefficient β was set to the best coefficient found in III.2. at 0.00015. After testing, two coefficients of β = 0.0003 and β = 0.002 were chosen to investigate the impact on London Underground flows when transport costs increase significantly.

To measure the impact of different transport cost functions on the flow in the network, since the departure and arrival flows are constant for all stations, we use the sum of the distances of all flows in the network, which represents the total number of passenger miles run on the London Underground, to measure the change in the operation of the network. We define the total flow distance in the network as:

$$\text{Total Flow Distance} = \sum_i \sum_j T_{ij} \times d_{ij}$$

Table 11 shows the performance of total flow distance for different cost function parameters β. We find that doubling β from 0.00015 to 0.0003 leads to a 24% reduction in the total flow distance, while when β is increased to 0.002, the flow of the network is reduced by 54%. This means that the actual number of passenger miles traveled on the London Underground has been reduced by more than half because passengers are influenced by transport costs to travel shorter distances.

Table 11: Total Flow Distance with Different Cost Function Coefficients

	Initial Flow	β = 0.00015	β = 0.0003	β = 0.002
Total Flow Distance	13,271,389,281	13,247,206,471	10,103,216,300	6,110,226,674
Reduction Percentage	0%	-0.2%	-23.9%	-54.0%

After β is greater than 0.002, the marginal effect on the flow decreases as β increases. This is since after the transport cost has increased to a certain level, there is no more room for variation after the destination of the flow at the departure point has been shifted to nearby stations. Therefore, when the transport cost increases to a certain level some more advanced spatial interaction model with the unlimited total number of flows should be considered.

Figure 4 shows the destinations of Waterloo, the station with the largest outbound flow under different traffic costs. Nearly a quarter of the outbound flow from Waterloo goes to Bank and Monument, but as transport costs increase, the flow to Bank and Monument decreases to 8% and 4%. With the β of 0.002, the head destinations from Waterloo become Westminster, Southwark, and Embankment, which are only one stop from Waterloo.

Figure 4: Outbound Flow from Waterloo Station

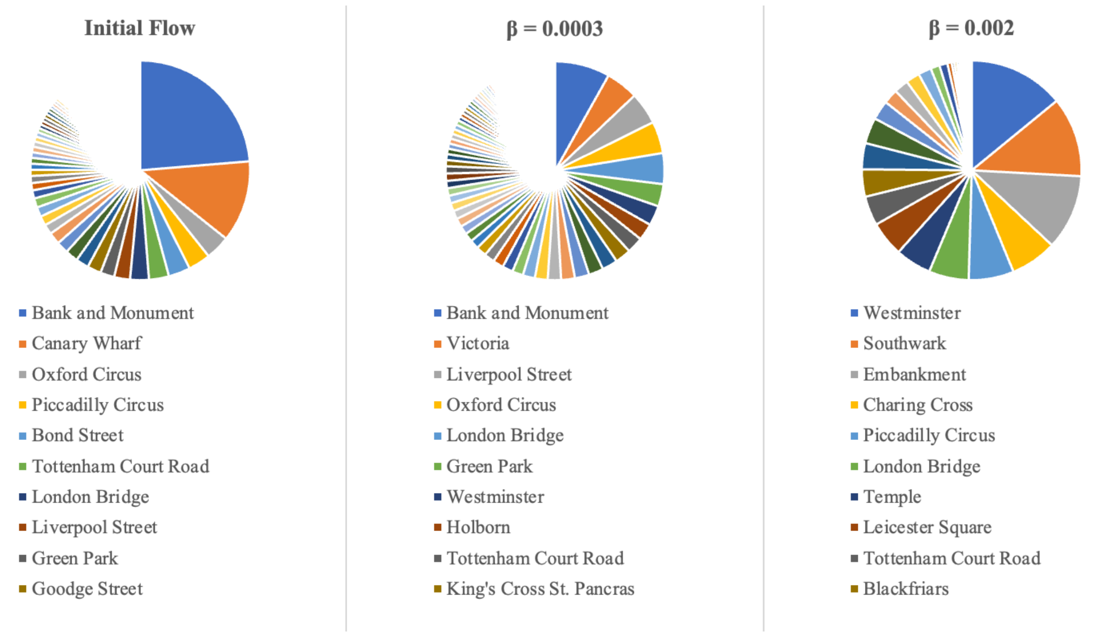
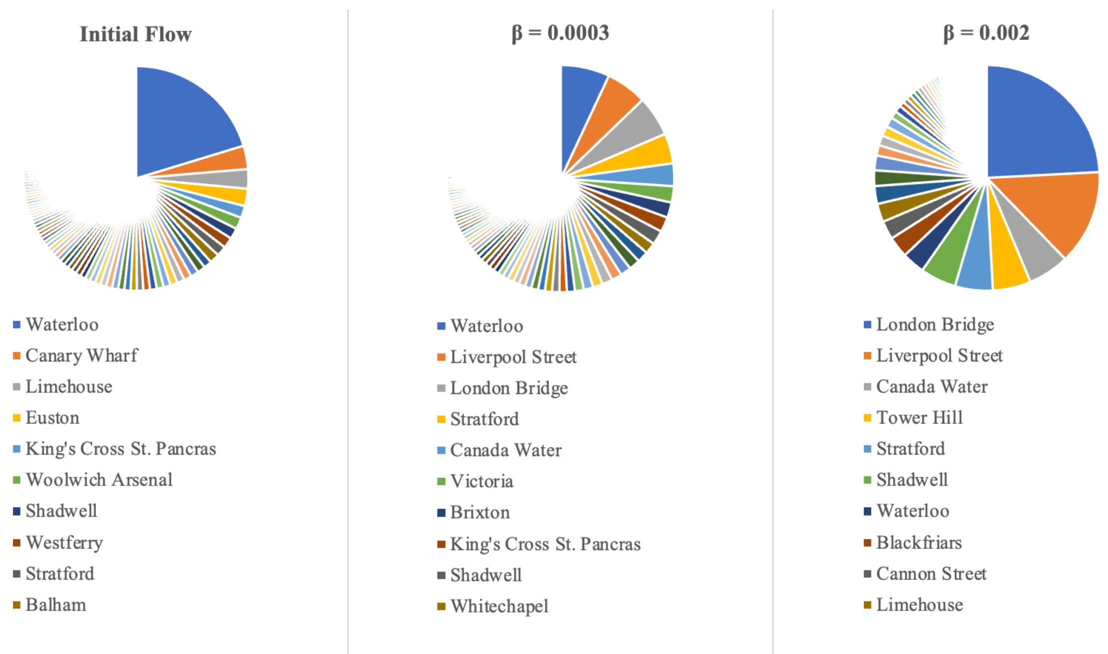


Figure 5 shows the variation in the sources of inbound flow at Bank and Monument, the station with the largest inbound flow, and we find similar patterns.

Figure 5: Inbound Flow to Bank and Monument Station



IV.3.

In Scenario A, the halving of jobs at Canary Wharf because of Brexit has had an impact on underground flows. Our model shows a 38% drop in flows to Canary Wharf across the Tube network. As a result of job losses, Canary Wharf has become less attractive to these locations and flows have shifted to other destinations. The largest increase in inbound flow is to Bank and Monument. From Figure 3 we can see that the main impact of the reduction in jobs in Canary Wharf is concentrated in the area around and along the tube through Canary Wharf, while the impact is moderate in other parts of London.

In Scenario B, we simulate the effect of increasing transport costs by adjusting the parameter of the cost function. We find that when transport costs increase, the effective distance carried by the tube system decreases significantly due to passengers traveling shorter distances, with the total flow distance decreasing by 24% when $\beta = 0.0003$ and by 54% when $\beta = 0.002$. The greater the β, the greater the impact on the underground network. Passengers' departure destinations will tend to be closer to their origin stations, and the source of flow to their destinations will shift to nearby stations, which may imply that the underground is underutilized.

For comparison, we define total flow change as the sum of the difference between the actual underground flow and the simulated flow.

$$\text{Total Flow Change} = \sum_{i}\sum_{j}\left|\widehat{T_{ij}} - T_{ij}\right|$$

Table 12 shows the total flow change for 3 different scenarios, and we find that the reduction in jobs at Canary Wharf has a similar flow redistribution effect to the model with $\beta = 0.0003$ rising transport costs, while the model with $\beta = 0.002$ had the most significant effect on flow redistribution.

Table 12: Total Flow Change in Different Scenarios

	Scenario A	**Scenario B**	
		β = 0.0003	**β = 0.002**
Total Flow Change	1,237,348	1,294,069	2,191,395

One possible reason is that the shocks caused in Scenario A are still geographical, and although there are significant impacts on Canary Wharf and the surrounding and connecting areas, the effects are less significant than those caused by transport costs, which have a direct impact on all flows in the network. Therefore, the overall transport cost of the network has a greater impact on the flow than the specific case of a major station in the network, so we think scenario B has a greater impact on the flow redistribution.

Reference

Wilson, A.G., 1971. A Family of Spatial Interaction Models, and Associated Developments. Environ Plan A 3, 1–32. https://doi.org/10.1068/a030001

Latora, V., Marchiori, M., 2001. Efficient Behavior of Small-World Networks. Phys. Rev. Lett. 87, 198701. https://doi.org/10.1103/PhysRevLett.87.198701

Kaiser, M., 2008. Mean clustering coefficients: the role of isolated nodes and leafs on clustering measures for small-world networks. New J. Phys. 10, 083042. https://doi.org/10.1088/1367-2630/10/8/083042

Senior, M.L., 1979. From gravity modelling to entropy maximizing: a pedagogic guide. Progress in Human Geography 3, 175–210. https://doi.org/10.1177/030913257900300218

NetworkX (no date) *Degree_centrality*, *degree_centrality - NetworkX 3.1 documentation*. Available at: https://networkx.org/documentation/stable/reference/algorithms/generated/networkx.algorithms.centrality.degree_centrality.html#networkx.algorithms.centrality.degree_centrality.

NetworkX (no date) *Closeness_centrality*, *closeness_centrality - NetworkX 3.1 documentation*. Available at: https://networkx.org/documentation/stable/reference/algorithms/generated/networkx.algorithms.centrality.closeness_centrality.html#networkx.algorithms.centrality.closeness_centrality.

NetworkX (no date) *Betweenness_centrality*, *betweenness_centrality - NetworkX 3.1 documentation*. Available at: https://networkx.org/documentation/stable/reference/algorithms/generated/networkx.algorithms.centrality.betweenness_centrality.html#networkx.algorithms.centrality.betweenness_centrality.

NetworkX (no date) *Transitivity*, *transitivity - NetworkX 3.1 documentation*. Available at: https://networkx.org/documentation/stable/reference/algorithms/generated/networkx.algorithms.cluster.transitivity.html#networkx.algorithms.cluster.transitivity

NetworkX (no date) *Average_clustering*, *average_clustering - NetworkX 3.1 documentation*. Available at: https://networkx.org/documentation/stable/reference/algorithms/generated/networkx.algorithms.cluster.average_clustering.html#networkx.algorithms.cluster.average_clustering.

NetworkX (no date) *Global_efficiency, global_efficiency - NetworkX 3.1 documentation*. Available at: https://networkx.org/documentation/stable/reference/algorithms/generated/networkx.algorithms.efficiency_measures.global_efficiency.html#networkx.algorithms.efficiency_measures.global_efficiency.

NetworkX (no date) *Average_shortest_path_length, average_shortest_path_length - NetworkX 3.1 documentation*. Available at: https://networkx.org/documentation/stable/reference/algorithms/generated/networkx.algorithms.shortest_paths.generic.average_shortest_path_length.html.