The code for each figure can be found in the corresponding folder. To generate each figure run the Matlab code with the corresponding name. These programs rely on subprograms; below I provide the description of these subprograms.

Common subprograms:

procedure.m: the key subprogram, which solves the allocation problem (3) from the paper. It takes cost draws "costs", the value of trade costs T, and iceberg trade costs option (1 for iceberg trade costs, 0 for specific trade costs) and returns the optimal allocation and the value of the value function.

fo_parents.m calculates the main clustering measure for given tree parameters. It uses procedure.m to calculate the optimal path and then checks if every pair of adjacent nodes share production location and calculates the share of such nodes for each pair of adjacent stages.

multiplier.m: for each downstream node there are M adjacent upstream neighbors; this subprogram is used by fo_parents.m (and similar programs such as fo_nieces.m) and multiplies downstream node M times to have 1-1 mapping of up- and downstream nodes.

progressbar.m is just a nice visualizer of the code progress, it is not essential for the programs to work.

Specific subprograms:

procedure_N1.m (used for Figures 7, 8, 12, A5, A6, A7): regular function procedure.m solves the allocation problem for an arbitrary complete tree but fails to work for the case of the snake (M=1) for syntax reasons. This program solves the allocation problem for M=1 where it is necessary.

fo_parents_N1.m (used for Figures 7 and 8): is a version of fo_parents.m, which works for the snake case (M=1).

fo_sisters.m, fo_nieces.m, fo_cousins.m, fo_grandparents.m, fo_greatgrandparents.m (used for Figures 11, A4): analogs of fo_parents but calculates the share of colocated nodes with sister, niece-aunt, cousins, grandparents-grandchildren, and greatgrandparents-greatgrandchildren relationships correspondingly.

f_reshoring_upper.m, f_reshoring_upper_N1.m (used for Figure 12, A5) calculate reshoring intensity for complete trees and snakes correspondingly. They run procedure.m for every value of trade costs, then calculate the share of nodes that were produced in the same location for high and low values of trade costs and elsewhere in between.

f_elasticity.m, f_elasticity_N1.m (used for Figures A6, A7, A8, A9, A10, A11) are conceptually similar to procedure.m and procedure_N1.m but their output is trade volumes at every production stage.

f_Cstatics.m (used for Figure 5) takes a given tree and solves the problem with the DP and LP methods and reports time that each of the methods took to solve the problem.

f_adjacency_direction (used for Figure 5): LP formulation follows the different logic and operates with a square matrix of node pairs rather than trees like the DP formulation. This code determines which node pairs in the LP formulation are adjacent and what their up-downstream relation is.

procedure_MC.m (used in Figure 6) is a version of procedure.m that reports total marginal costs of a firm rather than its optimal path.

Notes:

In the code M denotes the length of the tree and N is its length because it was the original notation in the earlier version of the paper.

Figure 5.m requires 16 GB of operative memory. If Matlab runs out of memory, try closing all background applications. On computers with more operative memories, the code can be run for larger values of N and M.

The code has been run on a Windows 10 computer with 16 GB of RAM and Intel(R) Core(TM) i7-10700 CPU @ 2.90GHz processor. I used Matlab R2021a. All of these programs except for Figure 5 were also run on Macbook 2020 Big Sur with 32 GB of RAM and 2.3 Ghz Quad-Core Interl Core i7 with Matlab R2021a.

Runtime:

Program	Runtime
Figure 5	160 hours
Figure 6	1.5 hours
Figure 7	1 hour
Figure 8	1 hour
Figure 9	1 hour 45 minutes
Figure 10	2 hours 45 minutes
Figure 11	3 hours 45 minutes
Figure 12	24 hours
Figure A2	1 hour 45 minutes
Figure A3	2 hours 45 minutes
Figure A4	4 hours 50 minutes
Figure A5	26 hours
Figure A6	7 hours
Figure A7	7 hours
Figure A8	4 hours 20 minutes
Figure A9	4 hours 20 minutes
Figure A10	6 hours 40 minutes
Figure A11	6 hours 40 minutes

12/11/2021