Folders and contents

- *1_AIS_summstats*
  - *code*:
    - *X.do*: cleans the raw AIS data and combines it with ship characteristics. Produces intermediate data sets used in steps 2-4 and stores them *in data_intermediate*
    - *01_dataprep_ais.do:* generates balanced port-to-port sample as input for the fastest routes calculation. Produces intermediate data sets used in step 2 and stores them in *data_intermediate*
    - *PortClusters.m* identifies clusters of ports that are close to each other.
    - *02_PanamaRoutes.do:* identifies direct port-to-port connections that pass the Panama Canal
    - *Y.do*: produces summary statistics on the AIS data
    - *z1-get-ship_all_variables.do:* clean raw AIS data and combine it with ship characteristics from clarkson *(these confidential original data files are not included in the replication package)*. Produce ship_all_variables.dta and intermdiate data for get ais_container_cleaned.dta and ship_port_to_port.dta2.
    - *z2-get-ais_container_cleaned.do:* produce ais_container_cleaned.dta, run after *z1-get-ship_all_variables.do*
    - *z3-get-ship_port_to_port.do:* produce ship_port_to_port.dta and erase intermediate data generated in z1-z3, run after *z2-get-ais_container_cleaned.do*
    - *Summstats-Tables_1_2_9.do:* code to produce summary statistics reported in Tables 1, 2, and 9.
      - *Output: .\1_AIS_data_summstats\output\T1-, T2- and T9-...xls*
- 2_fastest_routes
  - *code:*
    - *routes.do*: prepares the input for the fastest route algorithm and processes the output of the algorighm
    - *routes_hrs4saga.do*: runs the algorithm (requires array computing facility)
    - *descriptives_network*: produces figures and statistics describing the network
    - *routes_woPA*: prepares the input for the fastest route algorithm applied to the computation of fastest routes avoiding the Panama Canal and processes the output
    - *routes_hrs4sagaNP.do*: runs the algorithm for the fastest routes without the Panama Canal

        * *correlations_with_freight_cost.do*: compares fastest route travel times to freight cost inferred from US import data

- *3_DID*
  - *code:*
    * *dataprep_comtrade.do:* prepares the comtrade data
    * *exposure_measures.do:* prepares the main exposure measure and the alternative ones used in robustness checks
    * *DID_robustness.do*: prepares additional variables used in robustness checks
    * *DID.do:* runs the DID regression
- *4_counterfactuals*
  - code:
    * model.m: performs all counterfactuals
    * AIS_fixedpoint_hat.m, AIS_fixedpoint_hat_nonetwork.m : function to solve fixed points.
    * export_to_matlab.do: export data to csv files used in Matlab.
    * Get_port-pairs-data2.do : Gather data for all port pairs.
    * 10_expenditure by port.do : Gather data for expenditure by port.
    * margins.do : Results for "The Margins of Shipping" table.
- data_input:
  - BACI/ and Comtrade_agg: COMTRADE/BACII data
  - gravdataFTP1319quarterly.dta: WTO RTA database
  - Eora26_2015_pp/: Eora Global Supply Chain database
  - WDI/: World Development Indicators and INSEE

Software

- Stata 18.0
- Matlab R2019.b

Proprietary data

- The AIS port call data can be purchased from https://www.kpler.com/company/contact-us (formerly Marine Traffic).
- The Clarkson data can be purchased from research.crs@clarksons.com.