**0 Data Importation, Merging, and Collating**

Text in the published paper:

|  |
| --- |
| We obtain monthly flight data from the United States to all international destinations by US carriers for roughly the past three decades, from the [Bureau of Transportation Statistics](https://www.bts.gov/browse-statistical-products-and-data/bts-publications/data-bank-28im-t-100-and-t-100f-internationa-0). We also obtain annual macroeconomic data, including foreign direct investment estimates, for the same time period, from [World Bank Development Indicators](https://databank.worldbank.org/source/world-development-indicators). Finally, to focus on the impact of flight resumptions on FDI in conflict areas, we obtain daily [ACLED conflict data](https://acleddata.com/data-export-tool/) and aggregate to the annual level for each country.  The choice to focus on US carriers was not made lightly. There is passenger data available from the World Bank describing not just US carriers, but carriers around the world. But this data proved unobtainable or unsuited to our analysis. The World Bank posts the development indicator [“Air transport, passengers carried”](https://databank.worldbank.org/metadataglossary/world-development-indicators/series/IS.AIR.PSGR) which was defined as the domestic and international aircraft passengers of air carriers registered in the country in that year. This would be unsuitable to our analysis. For instance, if we wanted to see the associated FDI impacts of flights going to Jamaica, most if not all international flights would be on US carriers and would show up in the US’s data, not Jamaica’s. Plus, ratios of domestic to international travel for small countries’ national carriers would differ significantly between countries, so interpretation would vary significantly between countries. The metadata notes also mention that there may not be much international comparability because much is self-reported by the countries themselves to the ICAO and collection practice has not been “internationalized,” plus “for countries with few air carriers or only one, the addition or discontinuation of a home-based air carrier may cause significant changes in air traffic.” Therefore, this indicator will not work. We also found the [Global Aviation dashboard](https://www.worldbank.org/en/programs/global-aviation-dashboard) made by the World Bank, but downloads are not enabled. The authors plan to write to the World Bank to inquire about obtaining this data, which would be an improvement from a US-carrier-only passenger count. However, until this is achieved, the US passenger data appears to be consistent and complete, so we can make concrete inferences about the relationship between US carrier passengers flying into each country and changes in FDI inflows.  *~~Potential footnote: Regarding the “change in passengers flying into the country” variable, many outliers of >2500% growth (which is basically impossible without some data error) were produced by going from 0 to a high number of passengers. I chose to deal with this by dropping all observations on and before that year for such a country. (If it wasn't an extreme result, we keep it, because it still represents an increase from 0 which is meaningful. When calculating these cases, I divide the absolute change by a baseline of 100 passengers rather than 0. This absorbs some extreme results, but not all, which is why we need to handle the very extreme results by dropping them). Please view the Rmd for more details (starts at line 510).~~* |

Replication notes:

* I did not include some of the original data files, but they are downloadable from the website as specified in the Rmd. I also included code chunks to illustrate the processing from the point of download. These chunks are switched off so that once the file is run, there are no snags.
* Some setwd() may need to be changed to accommodate different local paths.
* full\_data.csv (stored in the “Replication Files” folder) is the output file.