\* This Do File finalizes the samples used in the empirical analyses

\* Set directory

cd "FILEPATH\Datasets"

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

\* FINALIZE SUBSIDIARY DATASET

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

\* Set dataset

use Subsidiary.dta, replace

rename \*, lower

\* Export Number of subsidiary-countries to MNC Level Dataset as control variable

capture frame drop FirmLevel

frame copy default FirmLevel

frame FirmLevel {

sort gvkey year

by gvkey year: egen nSubsidiary = count(size\_sub)

keep gvkey year nSubsidiary

duplicates drop gvkey year, force

save "nSubsidiary", replace

}

/\*\*\*\*\* Drop observations in countries where component auditor use does not vary to keep sample size consistent across countries \*\*\*\*\*/

quietly logit match size\_sub i.countryid

gen sample = e(sample)

drop if sample != 1

\* This results in the Final Sample used in the Subsidiary-Level Determinants and Audit Quality Analyses

\* The sample varies in size depending on the dependent variable and availability of control variables (see manuscript for details)

save Subsidiary\_final, replace

\*\*\*\* Aggregate Subsidiary Data for use in MNC Coverage Analyses

\* load dataset

cd "FILEPATH\Datasets"

use Subsidiary\_final, replace

gen obs = 1

\* Identify subsidiary sales audited by a component auditor

gen sale\_componentauditor = match \* sales

\* Sum totals by MNC parent company and year

sort gvkey year

\* Calculate total number of subsidiaries by MNC-Year

by gvkey year: egen SumSubs = sum(obs)

\* Calculate total subsidiary sales by MNC-Year

by gvkey year: egen total\_sub\_sale = sum(sales)

\* Calculate total sales at subsidiaries audited by a component auditor by MNC-Year

by gvkey year: egen total\_sale\_componentauditor = sum(sale\_componentauditor)

\* Calculate "Foreign Sales Coverage" - percentage of total subsidiary sales audited by component auditors

gen sales\_coverage = total\_sale\_componentauditor / total\_sub\_sale

\* Export Coverage variables for use in MNC Analyses

duplicates drop gvkey year, force

keep gvkey year SumSubs total\_sub\_sale total\_sale\_componentauditor sales\_coverage

save Coverage, replace

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

\*\*\*\* FINALIZE MNC DATASET

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

\*\*\*\*\*\* load dataset \*\*\*\*\*\*\*

cd "FILEPATH\Datasets"

use MNC, replace

rename \*, lower

\*\* Import number of subsidiary countries from subsidiary file

duplicates drop gvkey year, force

merge 1:1 gvkey year using "nSubsidiary"

drop \_merge

replace nSubsidiary = 0 if nSubsidiary == .

\* Calculate "MNC Ln Subsidiaries"

gen lnSubsidiary = log(nSubsidiary + 1)

\*\*\*\*\*\*\*\*\*\*\*\* Add Subsidiary Data from Orbis and Calculate Percentages \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

duplicates drop gvkey year, force

merge 1:1 gvkey year using "Coverage"

\* Calculate subsidiary sales as a percentage of MNC parent sales (used to restrict sample in Columns 1-2 of Table 6)

gen sub\_saleperc = total\_sub\_sale / sale

replace sub\_saleperc = 0 if sub\_saleperc < 0 & !missing(sub\_saleperc)

replace sub\_saleperc = 1 if sub\_saleperc > 1 & !missing(sub\_saleperc)

\* Generate variable equal to 1 if there is at least 1 Orbis subsidiary

gen Orbis = 0

replace Orbis = 1 if !missing(SumSubs)

\* Calculate Component Auditor "Any Coverage" - At least 1 Orbis Subsidiary and at least 1 Component auditor

gen coverage = Orbis == 1 & component == 1

replace sales\_coverage = 0 if missing(sales\_coverage)

\* This results in the Final Sample used in the MNC Audit Quality Analyses

save MNC\_final, replace