Total_sales_var_defn

- Sample: all assignments included in the USPTO assignment data. The sample period is labeled Jan 1970-Dec 2021 on the website. However, that dataset doesn't appear well populated until the 1980s and it is best practice to cut off the last few years of the sample period to make sure the transactions have had a chance to be recorded.
- Unit of observation: transaction id – firm buyer – firm seller

rf_id	Unique transaction ID per patent assignment data.
exec_dt_max	The latest execution date associated with an rf_id. Execution dates are defined for each individual assignor (seller), so if multiple sellers sign their agreement on a different date, there could be multiple execution dates for a given transaction.
exec_dt_min	The earliest execution date associated with an rf_id.
sale_year_max	The calendar year for exec_dt_max
sale_year_min	The calendar year for exec_dt_min
page_count	Page count of assignment record
lemon_sum	The total number of "lemons" included in a transaction. A lemon is defined as an application without a grant as identified in the assignment data.
lemon_avg	Averages of the underlying indicator variables above across patents in a transaction. A % of the transaction meeting the criteria.
app_sale_sum	The total number of patent applications included in a transaction. A patent is identified as an application at the time of the sale if it has non-missing date fields in the assignment data and exec_dt_min is before the grant date.
app_sale_avg	Averages of the underlying indicator variables above across patents in a transaction. A % of the transaction meeting the criteria.
grant_sale_sum	The total number of granted patents included in a transaction. A patent is identified as a grant at the time of the sale if it has non-missing date fields in the assignment data and exec_dt_min is after the grant date.
grant_sale_avg	Averages of the underlying indicator variables above across patents in a transaction. A % of the transaction meeting the criteria.
age_app_avg	The average age of a patent application in months (averaged across patents included in a transaction). Application age is defined the exec_dt_min less application date. This is negative in about 25% of the cases. My guess is this is due to applications that are part of a family and that the execution date is for an agreement that required later applications in the same family to be transferred as well. Age for a patent is missing if date values are missing in assignment data.
age_grant_avg	The average age of a granted patent in months (averaged across patents included in a transaction). Application age is defined the exec_dt_min less grant date. Age for a patent is missing if date values are missing in assignment data.
tot_pats	The total number of patents (either patent application or grant) with the same transaction id (rf_id)
no_seller_all	The total number of sellers for a given transaction, including any non-firm entities such as individuals or governments.
no_buyer_all	The total number of buyers for a given transaction, including any non-firm entities such as individuals or governments.
permno_seller	The permno associated with a seller in the transaction
permno_buyer	The permno associated with a buyer in the transaction

pub_seller	An indicator variable that takes on the value of one if the seller is a public firm and zero for private firms. The variable is missing if the seller is not a firm (e.g., an individual or government)
prv_seller	An indicator variable that takes on the value of one if the seller is a private firm and zero for public firms. The variable is missing if the seller is not a firm (e.g., an individual or government)
pub_buyer	An indicator variable that takes on the value of one if the buyer is a public firm and zero for private firms. The variable is missing if the buyer is not a firm (e.g., an individual or government)
prv_buyer	An indicator variable that takes on the value of one if the buyer is a private firm and zero for public firms. The variable is missing if the buyer is not a firm (e.g., an individual or government)
org_id_seller	An internal identifier created to link to the seller's organization type. It does not link to outside datasets.
org_id_buyer	An internal identifier created to link to the buyer's organization type. It does not link to outside datasets.
convey_ty	Conveyance type per the assignment dataset. Takes on the values assignment, govern or merger.
npe_buyer	An indicator variable that takes on the value of one if the buyer is an NPE and zero otherwise. The variable is missing if the buyer is not a firm (e.g., an individual or government)
npe_seller	An indicator variable that takes on the value of one if the seller is an NPE and zero otherwise. The variable is missing if the seller is not a firm (e.g., an individual or government)

Note: Variables tot_pats, page_count, no_seller, no_buyer, *_sum, and *_avg are all winsorized (see the 5-compile_dataset.do file row 231). So when merging any detailed patent-level information from the assignment data to our dataset, some observations may not agree with the total number of patents (tot_pats).