

## Cleaning data

As we can see, some of the continuous variables in the dataset has outliers. These outliers are located above the 99% percentile, and below the 1% percentile. To mitigate the influence of extreme values in the dataset, all continuous variables are therefore winsorized at these percentiles.

After checking for missing values, inconsistencies between private and public dummy values and negative values, I can't seem to find any mistakes in the data. The data fulfills the set of criteria defined in the assignment.

**Table 1**

Table 1 reports the annual averages of the listed variables, such as average deal value, average bidder CAR, share of deals with private targets and share of deals fully paid in stock.

Year	Avg. deal value	Avg. CAR	Private share	Stock share	Observations
1990	200.52	-0.01	0.448	0.478	67
1991	100.663	0.021	0.612	0.515	103
1992	108.129	0.024	0.719	0.597	139
1993	198.508	0.036	0.673	0.491	171
1994	278.832	0.018	0.59	0.556	234
1995	341.613	0.012	0.581	0.578	320
1996	404.285	0.022	0.58	0.589	338
1997	296.215	0.017	0.538	0.54	398
1998	556.577	0.006	0.5	0.5	430
1999	816.876	0.012	0.468	0.522	391
2000	992.903	-0.018	0.542	0.531	343
2001	736.371	-0.02	0.394	0.389	216
2002	327.898	-0.011	0.503	0.284	169
2003	401.478	0.015	0.485	0.263	171
2004	616.34	0	0.551	0.177	198
2005	952.07	-0.004	0.571	0.171	217
2006	833.448	0.006	0.51	0.161	192
2007	921.303	0	0.515	0.102	206
2008	758.906	0	0.598	0.165	127
2009	985.685	0.003	0.45	0.2	120
2010	925.069	0.008	0.437	0.118	119
2011	1182.967	-0.003	0.612	0.095	116
2012	939.275	0.029	0.491	0.069	116
2013	894.534	0.029	0.57	0.112	107
2014	1410.272	0.024	0.562	0.171	146

Table 1: Annual average statistics