6. Add control variables

6.1 Read data

In [1]:
```
cd ..
```

```
/home/julian/PycharmProjects/corporate_disruptions
```

In [2]:
```
import pandas as pd

company_year = pd.read_feather("preprocessed/company_year 2019-04-20.feather")
company_year.head()
```

Out[2]:

	gvkey	name	year	n_employees	n_exits	turnover	ceo_exit	recalls
0	025338	Bed Bath & Beyond Inc	2006.0	5	0	0.0	False	0.0
1	025338	Bed Bath & Beyond Inc	2007.0	5	0	0.0	False	0.0
2	025338	Bed Bath & Beyond Inc	2008.0	5	0	0.0	False	0.0
3	025338	Bed Bath & Beyond Inc	2009.0	5	0	0.0	False	0.0
4	025338	Bed Bath & Beyond Inc	2010.0	5	0	0.0	False	1.0

6.2 Connect to compustat

In [3]:
```
import parameters
import wrds_tools
```

In [4]:
```
wrds = wrds_tools.WrdsConnection(wrds_username=parameters.wrds_username)
```

```
Loading library list...
Done
```

In [5]:
```
from datetime import date

wrds.set_observation_period(start_date=date(year=2006, month=1, day=1),
                            end_date=date(year=2018, month=12, day=31))
```

In [6]:
```
wrds.dataset = company_year
```

6.3 Add revenue,

In [7]:
```
wrds.add_company_info(r_n_d=True, wages_and_salaries=True)
```

```
Warning: the data table you are about to download is very large. RAM and swap
partition usage might increase by 30GB or more. Press y to confirm.y
```

In [8]:
```
wrds.head()
```

Out[8]:

	gvkey	name	year	n_employees	n_exits	turnover	ceo_exit	recalls	revenue	r_n_d	wages_and
0	025338	Bed Bath & Beyond Inc	2006.0	5	0	0.0	False	0.0	6617.429	0.0	
1	025338	Bed Bath & Beyond Inc	2007.0	5	0	0.0	False	0.0	7048.942	0.0	
2	025338	Bed Bath & Beyond Inc	2008.0	5	0	0.0	False	0.0	7208.340	0.0	
3	025338	Bed Bath & Beyond Inc	2009.0	5	0	0.0	False	0.0	7828.793	0.0	
4	025338	Bed Bath & Beyond Inc	2010.0	5	0	0.0	False	1.0	8758.503	0.0	

In [9]:
```
with_controls = wrds.return_dataframe()
```

In [10]:
```
with_controls.to_feather('preprocessed/with_controls_{}.feather'.format(str(da
```

Make sure we have not missed any available data

In [11]:
```
test = '025338'
```

In [12]:
```
wrds.annuals_table[wrds.annuals_table['gvkey'] == test][['gvkey', 'revt', 'x
```

Out[12]:

	gvkey	revt	xrd	xstfws
475459	025338	167.595	0.0	NaN
475473	025338	216.712	0.0	NaN
475474	025338	305.767	0.0	NaN
475475	025338	440.261	0.0	NaN
475477	025338	601.252	0.0	NaN
475478	025338	823.178	0.0	NaN
475479	025338	NaN	NaN	NaN
475480	025338	1066.612	0.0	NaN
475481	025338	NaN	NaN	NaN
475482	025338	1397.197	0.0	NaN
475483	025338	NaN	NaN	NaN
475484	025338	1877.966	0.0	NaN
475485	025338	NaN	NaN	NaN
475487	025338	2396.655	0.0	NaN
475488	025338	NaN	NaN	NaN
475489	025338	2927.962	0.0	NaN
475490	025338	NaN	NaN	NaN
475492	025338	3665.164	0.0	NaN
475493	025338	NaN	NaN	NaN
475494	025338	4477.981	0.0	NaN
475495	025338	NaN	NaN	NaN
475497	025338	5147.678	0.0	NaN
475498	025338	NaN	NaN	NaN
475499	025338	5809.562	0.0	NaN
475500	025338	NaN	NaN	NaN
475502	025338	6617.429	0.0	NaN
475503	025338	NaN	NaN	NaN
475504	025338	7048.942	0.0	NaN
475505	025338	NaN	NaN	NaN
475507	025338	7208.340	0.0	NaN
475508	025338	NaN	NaN	NaN
475509	025338	7828.793	0.0	NaN