11_SBC_tables_sum_stat

April 17, 2020

1 Table 1 and 2

```
Service Google Drive and Docs, Sheet are now connected.
Service Google Drive is stored as <googleapiclient.discovery.Resource object at
0x1255d2a10> and accessible with "drive"
Service Google Doc is stored as <googleapiclient.discovery.Resource object at
0x125666c90> and accessible with "doc"
Service Google Sheet is stored as <googleapiclient.discovery.Resource object at
0x12574eb50>and accessible with "sheet"
Service account storage and Bigquery are now connected.
Service account storage is stored as <google.cloud.storage.client.Client object
at 0x12575a750> and accessible with "Storage_account"
Service account Bigquery is stored as <google.cloud.bigquery.client.Client
object at 0x12575a890> and accessible with "bigquery_account"
```

2 Load data

2.1 parameters

2.2 Load cityname_and_code from Google Spreadsheet

Feel free to add description about the dataset or any usefull information.

Profiling will be available soon for this dataset

```
  extra_coda geocode4_corr citycn   cityen citycn_correct cityen_correct  \
1       1100          1101        Beijing                       Beijing
2       1101          1101        Beijing                       Beijing
3       1102          1101        Beijing                       Beijing
4       1200          1201        Tianjin                       Tianjin
5       1201          1201        Tianjin                       Tianjin

  Province_cn Province_en
1               Beijing
2               Beijing
3               Beijing
4               Tianjin
5               Tianjin
```

2.3 Load TCZ_list_china from Google Spreadsheet

Feel free to add description about the dataset or any usefull information.

Profiling will be available soon for this dataset

	Province	City	geocode4_corr	TCZ
1	Beijing	Beijing	1101	1
2	Tianjin	Tianjin	1201	1
3	Hebei	Shijiazhuang	1301	1
4	Hebei	Tangshan	1302	1
5	Hebei	Qinhuangdao	1303	0

2.4 Load yearbook9813 from Google Spreadsheet

Feel free to add description about the dataset or any usefull information.

Profiling will be available soon for this dataset

We need this table for the table 0 -> Need year 2001

2.5 Paper dataset

2.6 Pollution China 1998-2005

2.7 Compute Herfhindal: proxy Size

$$H = \sum_{i=1}^{N} s_i^2$$

where s_i is the market share of industry[city] i in a city [industry], and N is the number of firms.

We proceed as follow: - Step 1: Compute the share [output, capital, employment] by city-industry: `market_share_cit` - Step 2: compute the sum of squared market share by industry[city]: `Herfindahl_agg_t` - Step 3: Compute the average across time: `Herfindahl_agg` - Step 4: Compute the deciles of step 3: `decile_herfhindal_agg` - Low decile implies a low concentration within sectors - High decile implies a high concentration within sectors

2.8 SOE vs Private

We proceed as follow: - Step 1: Compute the share [output, capital, employment] by industry[city], ownership (SOE/PRIVATE): `Share_X_agg_o` - ~Step 2: Compute dummy when share SOE above share domestic by industry[city]~ - Step 3: Compute decile by industry[city]-ownership - Note, high decile in SOE means the industry[city] has relatively high share of SOE output, but not in absolule value as in step 2. A decile 9 in SOE can be a decile 2 or 3 in Domestic

```
(228, 5)
```

3 Table 1

```
No SOE dominated    147
SOE dominated        81
Name: soe_city, dtype: int64

(228, 7)
```

3.0.1 1998-2000

- TCZ
- No TCZ
- SOE
- No SOE
- Full

3.0.2 2001-2005

- TCZ
- No TCZ
- SOE
- No SOE
- Full

3.0.3 2006-2010

Table 1: SO2 reduction during the subsequent FYPs

	1998-2001	2002-2005	2006-2010	Target
No TCZ	21.00%	64.00%	-11.00%	-6.00%
TCZ	-7.00%	38.00%	-15.00%	-16.00%
No Dominated SOE	-31.00%	30.00%	-16.00%	-10.00%
Dominated SOE	6.00%	21.00%	-17.00%	-12.00%
Full Sample	-2.00%	45.00%	-13.00%	-10.00%

Sources: Author's own computation The list of TCZ is provided by the State Council, 1998. "Official Reply to the State Council Concerning Acid Rain Control Areas and Sulfur Dioxide Pollution Control Areas". The information about the SO2 level are collected using various edition of the China Environment Statistics Yearbook. We compute the reduction of SO2 emission using the same methodology as Chen and al.(2018).

4 Table 2

Table 1: Summary statistics by city characteristics

index	Output share SOE_c	Capital share SOE_c	Employment share SOE_c
Full sample	24.95%	35.48%	29.19%
Central	28.78%	41.80%	34.17%
Coastal	11.28%	18.50%	12.04%
Northeast	26.16%	38.73%	33.83%
Northwest	43.38%	51.06%	49.53%
Southwest	31.02%	44.59%	36.08%
Central	28.78%	41.80%	34.17%
Eastern	15.99%	24.83%	19.05%
Western	37.71%	48.44%	42.82%
No TCZ	24.32%	35.07%	29.97%
TCZ	25.34%	35.74%	28.70%
Concentrated city	34.09%	45.39%	37.89%
No Concentrated city	15.14%	24.86%	19.86%

Sources: Author's own computation The list of TCZ is provided by the State Council, 1998. Output share SOE, Capital share SOE, Employment share SOE is computed using the average output, capital or employment share by city over 2002-2005

5 Create report