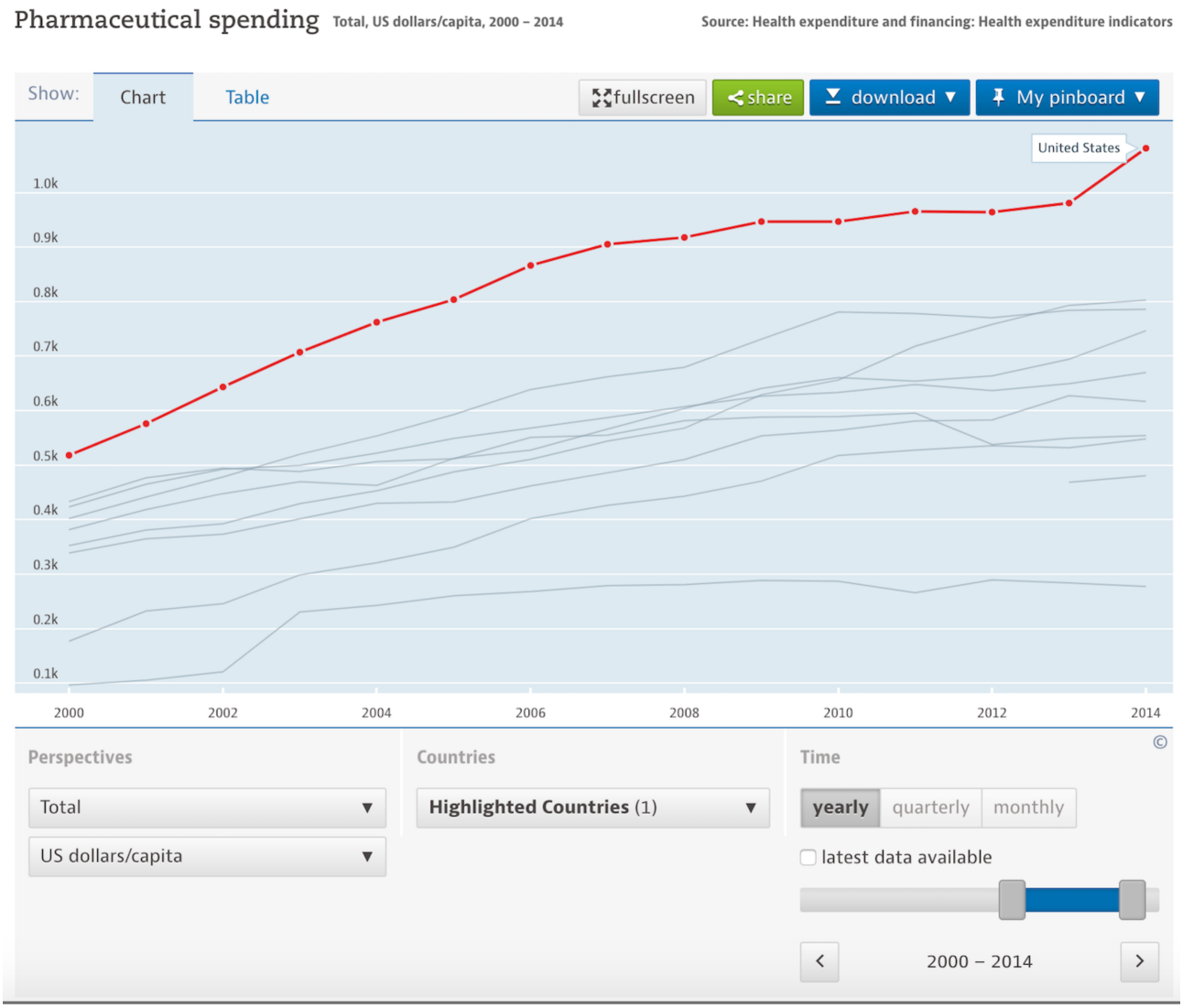
DSDHT FALL 2017
Carlos Sathler (cssathler@gmail.com)

Module 5 – Drug Discovery Economics; following the money

1. The Organization for Economic Cooperation and Development website has data relevant to health and pharma economics. Go to https://data.oecd.org/healthres/pharmaceutical-spending.htm where there is an interactive tool for exploring international pharma spending data. Use this tool to plot pharma spending per capita in the G20 countries from 2000-2014. Highlight the US plot and save this plot as an image file.



2. Use the tool to view measure "% of GDP", and then download the "Selected data only" CSV file from the OECD site (revised instructions, since previous version of the tool included this in "Full indicator data".) Using Tableau or another tool, plot health spending the percent of GDP (Measure = "PC_GDP") for Canada, France, Great Britain, Japan, Sweden and USA, from 2000-2014. Label the axes and save plot as an image file.

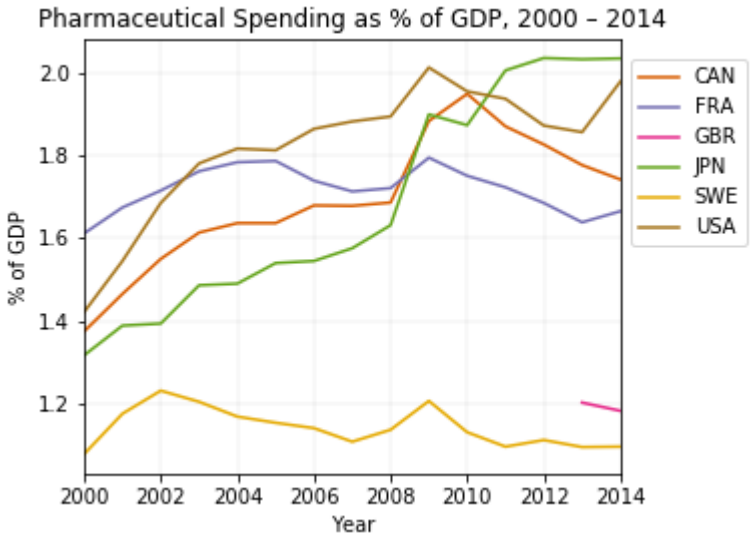


Figure 1

3. From these data, and the data presented in the lectures, what can you interpret about pharmaceutical spending? Is the pharmaceutical economic sector growing?

   Based on what we learned in class, particularly regarding R&D spending for drug discovery, there has been growth, as seen for example in the "Evolution of R&D Spending…" plot, reproduced here from the class slides. Also from the per capita spending from question 1 above.

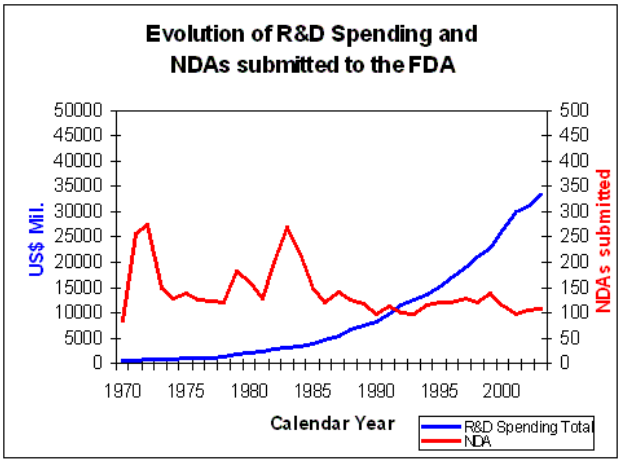


Figure 2 – Obtained from class slides

More importantly, the data on pharmaceutical spending in the OECD website, for years between 2000 and 2014 clearly shows that the average pharmaceutical spending across

all countries in the OECD dataset, measured as the average % of GDP, has clearly increased. That is true, on average, for Canada, France, Japan, Sweden, Great Britain and the USA, as well as all other countries, as seen in Figure 3 below. In fact, the first set of countries show an approximate increase of 16% (from about 1.375 to about 1.6 % of GDP, on average), which is similar to the increase for the remaining countries (from 1.2 to about 1.4 % of GDP, on average).

On a final note it is interesting to see that countries like Japan, Canada and the USA have shown higher spending (as % of GDP) in relation to others in the list of selected countries from question 2. Particularly Japan has shown a steady increase which is unmatched by other countries. Sweden and France, on the other hand, have not seen an increase in spending relative to their GDP. It is also interesting to notice a spike in spending in 2009, followed by a decline that only seem to have reverted in 2013 for the countries in the list selected from question 2.

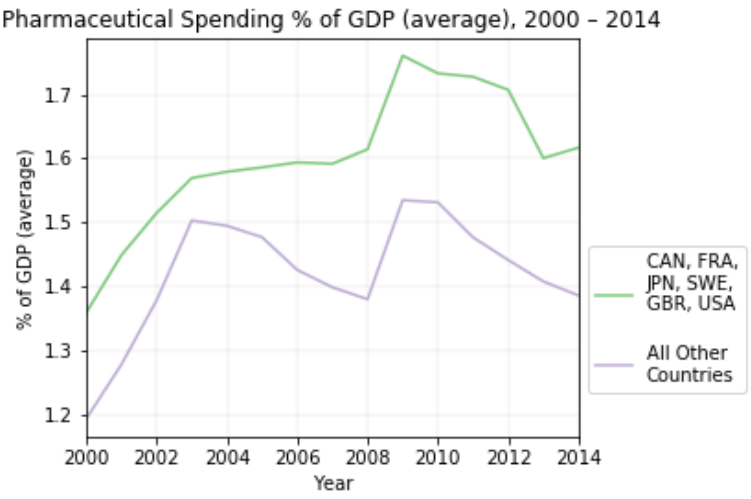


Figure 3

Note: For python code used to generate plots, please refer to Appendix

DSDHT FALL 2017
Carlos Sathler (cssathler@gmail.com)

Appendix A – Jupyter notebook exported as .py code

```
# coding: utf-8

# In[1]:
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from matplotlib import cm
get_ipython().magic('matplotlib inline')

# In[2]:
df = pd.read_csv('DP_LIVE_21092017145828500.csv')
df.head()

# In[3]:
df.LOCATION.unique()

# In[4]:
countries = ['CAN','FRA','GBR','JPN','SWE','USA']
df_countries = df[df.LOCATION.isin(countries)]
df_countries.LOCATION.unique()

# In[5]:
# Some of this code based on Stackoverflow entry for "Plotting multiple lines with
pandas dataframe"
# URL: https://stackoverflow.com/questions/29233283/plotting-multiple-lines-with-
pandas-dataframe
# Author: unutbu

fig, ax = plt.subplots()
labels = []
i = 1
for key, grp in df_countries.groupby(['LOCATION']):
    ax = grp.plot(ax=ax, kind='line', x='TIME', y='Value', c=cm.Dark2(i))
    labels.append(key)
    i += 1
lines, _ = ax.get_legend_handles_labels()

# Shrink current axis by 20%
# You could put the legend at top, botton, right, or left; you would shrink the
proper dimension to accomodate legend
box = ax.get_position()
ax.set_position([box.x0, box.y0, box.width * 0.8, box.height])
ax.set_title('Pharmaceutical Spending as % of GDP, 2000 ,Äì 2014')
ax.set_xlabel('Year')
ax.set_ylabel('% of GDP')
ax.grid(color='grey', linestyle='-', linewidth=0.1)

# Put a legend to the right of the current axis
ax.legend(lines, labels, loc='lower left', bbox_to_anchor=(1, 0.5))

plt.savefig('module5.png')

# In[6]:
# produce totals by year, for the selected countries from assigment
df_total_countries = df_countries.groupby('TIME').agg({'Value': np.mean})
df_total_countries = df_total_countries.reset_index()

# produce totals by year, for all other countries
```

```
df_other_countries = df[df.LOCATION.isin(countries)==False]
df_total_other_countries = df_other_countries.groupby('TIME').agg({'Value':
np.mean})
df_total_other_countries = df_total_other_countries.reset_index()

# In[7]:
# Some of this code based on Stackoverflow entry for "Plotting multiple lines with
pandas dataframe"
# URL: https://stackoverflow.com/questions/29233283/plotting-multiple-lines-with-
pandas-dataframe
# Author: unutbu

fig, ax = plt.subplots()
labels = []
ax = df_total_countries.plot(ax=ax, kind='line', x='TIME', y='Value',
c=cm.Accent(0))
labels.append('CAN, FRA,\nJPN, SWE,\nGBR, USA\n')
ax = df_total_other_countries.plot(ax=ax, kind='line', x='TIME', y='Value',
c=cm.Accent(1))
labels.append('All Other\nCountries')
lines, _ = ax.get_legend_handles_labels()

# Shrink current axis by 20%
# You could put the legend at top, botton, right, or left; you would shrink the
proper dimension to accomodate legend
box = ax.get_position()
ax.set_position([box.x0, box.y0, box.width * 0.8, box.height])
ax.set_title('Pharmaceutical Spending % of GDP (average), 2000 ,Äì 2014')
ax.set_xlabel('Year')
ax.set_ylabel('% of GDP (average)')
ax.grid(color='grey', linestyle='-', linewidth=0.1)

# Put a legend to the right of the current axis
ax.legend(lines, labels, loc='upper left', bbox_to_anchor=(1, 0.5))

plt.savefig('module5_2.png', orientation='landscape')
```