

US FEDERAL DEBT PROJECT

DOCUMENTATION : INDUSTRIAL SUITE

DRAFT

George Hall*, Jonathan Payne†, Tom Sargent‡

July 14, 2017

Contents

1	UNDERLYING EXCEL FILES	2
1.1	RAWDATA/BONDLISTDATA FOLDER:	2
1.1.1	Description of Bond_List_1776_1918	3
1.1.2	Description of Bond_List_1919_1949 and Bond_List_1950_1960	5
1.1.3	Bond Categorization	7
1.2	RAWDATA/BONDQUANT FOLDER:	10
1.2.1	Series for L1 ID 2 (Interest Bearing)	11
1.2.2	Series for L1 ID 1 (Pre 1790 Domestic Debt) and L1 ID 4 (Other)	13
1.2.3	Series for L1 ID 3 (Non-Interest Bearing)	13
1.2.4	Series for L1 ID 5 (Asset)	13
1.3	RAWDATA/BONDPRICE FOLDER:	13
1.4	RAWDATA/MACRO DATA FOLDER	15
1.5	RAWDATA/ISSUEREDEMPTIONS FOLDER	16
2	CODE FOR IMPORTING THE EXCEL FILES	16
2.1	NOTES ON IMPORTING THE LIST DATA	16
2.2	NOTES ON IMPORTING THE QUANTITY DATA	16
2.3	NOTES ON IMPORTING THE PRICE DATA	16
3	DATA FRAMES	18
3.1	STORAGE FORMATS	18
3.2	BONDLIST DATA FRAME	18
3.3	BONDQUANT DATA FRAME	20

*Brandeis University, Department of Economics. Email: ghall@brandeis.edu

†NYU, Department of Economics. Email: jep459@nyu.edu

‡NYU, Department of Economics. Email: thomas.sargent@nyu.edu

3.4 BONDPRICE DATA FRAME	21
------------------------------------	----

INTRODUCTION

This document describes the organization and interpretation of the files in the ‘Industrial Suite’ of the US Federal Debt Project. It should be read before running the code in any of the folders. The document has the following parts:

1. Underlying Excel Files: explains the location and organization of the excel files in which the data was originally recorded. These excels files are the most faithful recording of the primary sources but are also the least standardized.
2. Import Code Files: (briefly) explains the location of the python files that import the code from the underlying excel files and ‘standardize’ it into python data frames.
3. Data Frames: explains how to access and interpret the final python data frames that store the ‘standardized’ data.

1 UNDERLYING EXCEL FILES

The data was initially recorded into a collection of data files located in the `RawData` folder. There are five folders with different types of files:

1. `RawData/BondListData`: contains descriptive information for each individual Treasury issue
2. `RawData/BondQuantData`: contains time series information on the quantity outstanding for each issue
3. `RawData/BondPriceData`: contains time series information on the price of each issue
4. `RawData/IssuesRedData`: contains time series information on the issues and redemption made for each issue
5. `RawData/MacroData`: contains time series data on key macroeconomic variables.

1.1 RAWDATA/BONDLISTDATA FOLDER:

Descriptions of each issue are reported in the following files:

- `Bond_List_1776_1918.xlsx`
- `Bond_List_1919_1949.xlsx`
- `Bond_List_1950_1960.xlsx`

For securities issued prior to 1880, the description of each security is taken from Bayley [1882]. For securities issues after 1880, the descriptions are take from various issues of the United States Department of the Treasury [1869-2015].

The content of the different files are described in the subsubsections below.

1.1.1 Description of Bond_List_1776_1918

The column headings for the file `Bond_List_1776_1918.xlsx` (listed in the order that they appear in the excel file) have the following interpretations:

- *Sheet*: Describes which date sheet contains the quantity and price information for that bond:
 - 10000: securities issued between 1776 and 1840
 - 20000: securities issued between 1840 and 1900
 - 30000: securities issued between 1900 and 1918
- *Loan Number*: unique number assigned to each security by sheet
- *Continuation*: the sheet and loan number for the security in the prior sheet category (if such a security exists).
- *Treasury's Name of Issue*: Each security was given a name by the US Treasury. Often the name is based on the year that the security was issued. We refer to each security by its official name as recorded by Bayley [1882] or the Monthly Statement of the Public Debt.

The public often referred to securities using names based on the coupon rate and the maturity date. For example, the bonds of the “Loan of 1860” are reported in newspapers as the “U. S. Fives of 1871”.

- *Temp ID*: Contains the sum of Sheet and Loan
- *Import ID*: Contains the sum of Link-Sheet and Link-Loan if they are non-empty. Otherwise, it contains the sum of Sheet and Loan. The Import ID is the unique identifier for a particular bond. It is used to identify the bond in the price and quantity time series spreadsheets.
- *Bayley Page*: page number of Bayley [1882] which describes the security.
- *Authorizing Act*: The Act of Congress that authorized the Secretary of the Treasury to issue the security and set the terms of the security.
- *Date of First Issue*: End of the first month that the security appears on the Treasury's records.
- *Price sold/Par*: Ratio of the initial price of the security to its par value. A value less (more) than one denotes that the bond initially sold at a discount (premium).

- *Coin*: Value of 1 if payable in gold or silver; value of 0 if payable in lawful money.
- *Term*: The length (in years) of the loan.
- *First Redemption, Final Redemption, Redeemable After and Payable*: There are four different column categories to indicate the maturity of the bond:
 - If the security has a single fixed maturity date, then this date is recorded under *Payable*.
 - If the security is redeemable at the option of the Treasury after a single date, then this date is recorded under *Redeemable After*.
 - If the security does not have a fixed maturity or redemption date, or if the repayment of the security is scheduled across multiple dates, the first and last redemption dates are recorded under the *First Redemption* and *Final Redemption* columns.
- *Callable*: If the security is callable after a single date prior to a single maturity date, then there is a 1 in this column. Otherwise the value is set to 0. If the bond is callable, then the call date is recorded in the *Redeemable* column and the maturity date is recorded in the *Payable* column.
- *Coupon Rate*: The sum of the coupon or interest payments promised each year divided by the face value and expressed as a percent. If there is a single coupon rate, then this rate is recorded in the *Unique* column; if there is a range of coupon rates, then minimum and maximum of the range is reported in the *Min* and *Max* columns.
- *Coupon Frequency* and *Coupons per year*: The number of coupon payments per year.
- *Coupon Schedule*: Entries with 1 denote the months coupon payments are due. If the coupon schedule is ambiguous, there is a 1 in the *Ambiguous* column.
- *Statutory Authorization*:
 - The first column, *Amount Authorized*, reports the total dollar amount Congress authorized the Treasury to borrow.
 - The second column, *Limit On Issued*, records a 1 if the statutory limit is on quantity issued and a 0 if the statutory limit is on quantity outstanding.
 - The third column, *Marginal Increase in Total Debt Authorized* records the marginal increase to total quantity of debt outstanding.
- *Primary Loan Purpose*: Statutory purpose for the security
 - 1 to finance military spending

- 2 to finance civil spending
- 3 to refinance existing debt
- 4 to finance interest payments
- 5 to buy gold or silver to back currency
- 6 debt bearing no interest
- 7 certificates and notes issued on deposits
- 8 matured debt
- 9 national bank bonds
- 10 postal savings bonds
- 99 other
- *Categories*: The category columns contains the categorization of the bonds. (The categorization scheme is explained in further detail in the subsection 1.1.3.)
 - *Category L1*: gives the level 1 category of the bond (the highest level category of the bond)
 - *Category L2*: gives the level 2 category of the bond
 - *Category L3*: gives the level 3 category of the bond (the lowest level category of the bond)
- *Category IDs*: The category ID columns contain the numerical IDs corresponding to the categories. (The categorization scheme is explained in further detail in the subsection 1.1.3.)
- The remaining columns index the relevant quantity and price files.

1.1.2 Description of **Bond_List_1919_1949** and **Bond_List_1950_1960**

In addition to a subset of the columns described above, the **Bond_List_1919_1949.xlsx** spreadsheet contains two additional columns:

- *C.F. Childs Page*: page number of C.F. Childs [1947] which describes the security, and
- *Tax Status* : categories given by the US Treasury (and described below).

The **Bond_List_1950_1960.xlsx** spreadsheet contains the *Tax Status* column, but not the *C.F. Childs Page* column.

The tax status for each issue reported by the US Treasury as follows:

- d. Any income derived from Treasury bills of this issue, whether interest or gain from their sale or other disposition does not have any exemption, as such, and loss from the sale or other disposition of any such bills does not have any special treatment, as such, under the Internal Revenue Code

or laws amendatory or supplementary thereto. The bills are subject to estate, inheritance, gift, or other excise taxes, whether Federal or State, but are exempt from all taxation now or hereafter imposed on the principal or interest thereof by any State, or any of the possessions of the United States, or by any local taxing authority. For purposes of taxation the amount of discount at which the bills are originally sold by the United States is to be considered to be interest.

- e. Income derived from these securities is subject to all taxes now or hereafter imposed under the Internal Revenue Code or laws amendatory or supplementary thereto. The securities are subject to estate, inheritance, gift, or other excise taxes, whether Federal or State, but are exempt from all taxation now or hereafter imposed on the principal or interest thereof by any State, or any of the possessions of the United States, or by any local taxing authority. The following is applicable to savings bonds only: For the purposes of taxation any increment in value of savings bonds represented by the difference between the price paid and the redemption value received (whether at or before maturity) shall be considered as interest.

Attention is invited to Treasury Decision 4550 ruling that bonds, notes, bills, and certificates of indebtedness of the Federal Government or its agencies, and the interest thereon, are not exempt from the gift tax.

- f. Exempt both as to principal and interest, from all taxation now or hereafter imposed by the United States, any State, or any of the possessions of the United States, or by any local taxing authority, except (a) estate or inheritance taxes, and (b) graduated additional income taxes, commonly known as surtaxes, and excess profits and war-profits taxes, now or hereafter imposed by the United States, upon the income or profits of individuals, partnerships, associations, or corporations. The interest on an amount of bonds authorized by the act approved Sept. 24, 1917, as amended, the principal of which does not exceed in the aggregate \$5,000 owned by any individual, partnership, association, or corporation, shall be exempt from the taxes provided for in clause (b) above. The following is applicable to savings bonds only: For the purposes of determining taxes and tax exemptions the increment in value of savings bonds represented by the difference between the price paid and the redemption value received (whether at or before maturity) shall be considered as interest.
- g. Exempt from the payment of all taxes or duties of the United States as well as from all taxation in any form by or under State, municipal, or local authority. (The Supreme Court has held that this exemption does not extend to estate or inheritance taxes, imposed by Federal or State authority.)
- h. Interest on these bonds is subject to all taxes now or hereafter imposed under the Internal Revenue Code or laws amendatory or supplementary thereto. Principal is exempt from taxation.

- i. These issues being investments of various Government funds and payable only for the account of such funds have no present tax liability. In hands of foreign holders-Applicable only to securities issued prior to Mar. 1, 1941: Bonds, notes, and certificates of indebtedness of the United States, shall, while beneficially owned by a nonresident alien individual, or a foreign corporation, partnership, or association, not engaged in business in the United States, be exempt both as to principal and interest from any and all taxation now or hereafter imposed by the United States, any State, or any of the possessions of the United States, or by any local taxing authority.

1.1.3 Bond Categorization

The bonds have been categorized hierarchically into three levels. These levels are shown in table 1 and described in the list below:

- Category Level 1 (L1):
 1. *Pre 1790 Domestic Debt*: Debt issued before 1790 (mostly during the revolutionary war).
 2. *Interest Bearing*: Bonds that pay interest.
 3. *Non-Interest Bearing*: Bonds that do not pay interest.
 4. *Other*: Miscellaneous bonds and adjustments
 5. *Asset*: Assets held by treasury (incomplete)
- Category Level 2 (L2):
 1. *Pre 1790 Domestic Debt*: Same as L1
 2. *Marketable*: Interest bearing bonds that are publicly traded.
 3. *Non Marketable*: Interest bearing bonds that are not publicly traded
 4. *Non-Interest Bearing*: Same as L1
 5. *Other*: Same as L1
 6. *Asset*: Same as L1
- Category Level 3 (L3):
 1. *Loan Office Certificates*: Contains certificates sold to by the revolutionary government in order to fund the war of independence
 2. *Interest In Arrears*: Unpaid interest prior to the Funding Act of 1790.
 3. *Foreign Loan*: Contains loans made by foreign governments to the US government. These were mostly made in the period 1776-1840.
 4. *Temporary Loan*: Contains loans made to the US government before 1920 which the treasury called "temporary loans" at the date of issue. These loans typically have duration of 1-2 years.

Category L1	Category L2	Category L3
Pre 1790 Domestic Debt	Pre 1790 Domestic Debt	Loan Office Certificates
		Interest In Arrears
Interest Bearing	Marketable	Foreign Loan
		Temporary Loan
		Long Term Bond
		Treasury Note : Pre 1920
		Treasury Note : Post 1920
		Treasury Bond
		Treasury Bill
		Certificates Of Indebtedness
		Navy Pension Fund
		Panama Canal Bond
		Postal Savings Bond
		Liberty Loan
		War Savings Bond
	Marketable	Savings Notes
		Savings Bonds
		Armed Forces Leave Bonds
		Depository Bond
		Treasury Bonds Investment Series
		Special Issues
Non-Interest Bearing	Non-Interest Bearing	Commodity Currency
		Fiat Currency
Other	Other	Old Debt
		Pacific Railroad
		Old Loan
		Prepayments
		Adjustment
		Bounty Land Script
Asset	Asset	Asset

Table 1: Bond Categories

5. *Long Term Bond*: Contains all loans, consols, stocks, refunding certificates other long term bonds issued by the US government. They typically have duration greater than 10 years.
6. *Treasury Note : Pre 1920*: Contains all treasury notes issues by the US government before 1920. (The division is made because the US government debt market was "standardized" in 1920 and so Treasury Notes issued after that date are different in nature.)
7. *Treasury Note : Post 1920*: Contains all treasury notes issues by the US government after 1920. (The division is made because the US government debt market was "standardized" in 1920 and so Treasury Notes issued after that date are different in nature.)
8. *Treasury Bond*: Contains all treasury bonds issued by the US government. All treasury bonds were issued after the US government debt market was "standardized" in 1920.
9. *Treasury Bill*: Contains all treasury bills issued by the US government. All treasury bills were issued after the US government debt market was "standardized" in 1920.
10. *Certificates Of Indebtedness*: Contains certificates of indebtedness
11. *Navy Pension Fund*: Contains Navy pension funds.
12. *Panama Canal Bond*: Contains all bonds issue to fund the Panama Canal.
13. *Postal Savings Bond*: Contains bonds issued in lieu of postal savings deposits or certificates in denominations of \$20, \$100, and \$500. The issuance of Postal Savings Bonds was discontinued on July 1, 1935. U.S. Savings Bonds replaced Postal Savings Bonds.
14. *Liberty Loan*: Contains all Liberty Loans issued by the US government. These were bonds sold to fund to the first world war.
15. *War Savings Bond*: Contains savings bonds issued in order to fund the second world war.
16. *Savings Notes*: Contains all Savings Notes.
17. *Savings Bonds*: Contain all US Saving Bonds issued by the government (not including War Savings Bonds).
18. *Armed Forces Leave Bonds*: Contains bonds issued as compensation for accumulated leave to members and former members of the Armed Forces of World War II.
19. *Depository Bond*: Contains bonds held by banks to collateralise deposits.
20. *Treasury Bonds Investment Series*:
21. *Special Issues*: Includes a variety of items including: Adjusted Service Series, Adjusted Service Certificate Fund, Unemployment Trust

Fund, Fed Old-Age and Survivors Ins. Trust Fund Series, Government Life Insurance Fund Series, Civil Service Retirement Fund Series, Railroad Retirement Account, Foreign Service Retirement Fund Series, Canal Zone Retirement Fund Series, Alaska Railroad Retirement Fund Series, Postal Savings System, Gov.Life Insurance Fund Series, Federal S&L Insurance Corp Series and National Service Life Insurance Fund.

22. *Commodity Currency*: Currencies backed by a commodity. This includes Certificates of Deposit, Gold Certificates, Silver Certificates, Treasury Notes of 1890
23. *Fiat Currency*: Currencies not back by a commodity.
24. *Old Debt*: Matured issues still outstanding
25. *Pacific Railroad*: Bonds issued to subsidize the construction of the trans-continental railroads.
26. *Old Loan*: Includes the series Old Debt, Old Demand Notes and Carry-over from Previous Spreadsheet. These are series summarising (matured) US government liabilities continuing from previous spreadsheets.
27. *Prepayments*: **XXX**
28. *Adjustment*: **XXX**
29. *Bounty Land Script*: Land warrants issued as payment for military service.
30. *Asset*: **XXX**

1.2 RAWDATA/BONDQUANT FOLDER:

The quantity time series data for the bonds is spread across the following files:

- Gov_Bond_Quant_1776_1840.xlsx
- Gov_Bond_Quant_1840_1900.xlsx
- Gov_Bond_Quant_1900_1918.xlsx
- Gov_Bond_Quant_1919_1949.xlsx
- Gov_Bond_Quant_1950_1960.xlsx
- Gov_Bond_Gov_Accounts_1919_1949.xlsx
- Gov_Bond_Gov_Accounts_1950_1960.xlsx

The quantities outstanding from 1790 to 1871 are imputed from the issue and redemption series reported by Bayley [1882]; These quantities were cross-checked with with quantity outstanding series reported in Register's Office [1886]. After

1871 the source of quantity outstanding series is the United States Department of the Treasury [1869-2015] *Monthly Statements of the Public Debt*.

The call data are from various years of the *Annual Reports of the Secretary of Treasury*.

Data on Treasury securities held in government accounts are from *Banking and Monetary Statistics 1914-1941* prior to 1941 and from *Treasury Bulletin* thereafter.¹

Depending on its category, a particular bond may have multiple different time series. In the following subsections, the different possible series types are described for the following categories:

- L1 ID 2 (Interest Bearing)
- L1 ID 1 (Pre 1790 Domestic Debt) and L1 ID 4 (Other)
- L1 ID 3 (Non-Interest Bearing)
- L1 ID 5 (Asset)

1.2.1 Series for L1 ID 2 (Interest Bearing)

Interest Bearing bonds have the largest number of possible series. Figure 1 shows, diagrammatically, how all the different possible types of series fit together (although not all bonds will have all of these series).

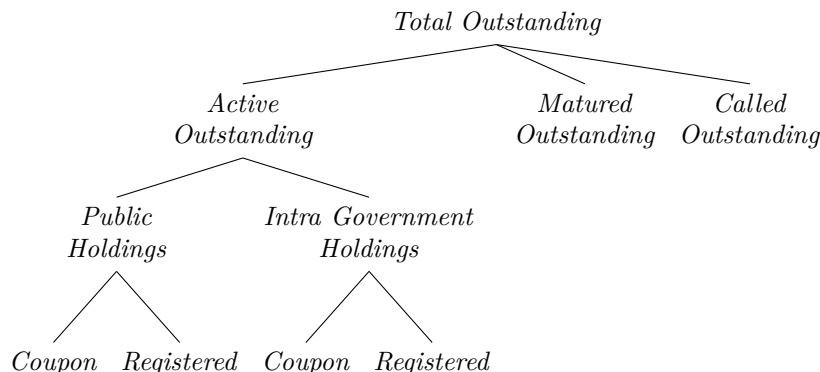


Figure 1: Types of Quantity Series: Interest Bearing Bonds

As can be seen, there are three levels of tree branching:

- Level 1: The *Total Outstanding* quantity of a particular bond is divided into the following categories:

¹See Board of Governors of the Federal Reserve System [1943] and United States Department of the Treasury [1939-2015].

- *Active Outstanding*: Contains the quantity that has not yet been redeemed (either because the maturity has passed or the security has been called)
- *Matured Outstanding*: Contains the quantity that has matured but has not yet been collected
- *Called Outstanding*: Contains the quantity that has been called but has not yet been collected
- Level 2: *Active Outstanding*: is divided into the following categories:
 - *Public Holdings*: Contains the quantity held by the public
 - *Intra Government Holdings*: Contains the quantity held by the Government. The data set only contains information on Intra Government Holdings after the period 1919.
- Level 3: *Public Holdings* and *Intra Government Holdings* are divided up into the following categories:
 - *Coupon*: contains the sum of all active liabilities outstanding that gave bond holders a physical coupon book.
 - *Registered*: contains the sum of all non-matured liabilities outstanding that registered scheduled coupon payments in a government ledger.

XXXXX JP: George, in principle, should there only be Coupon and Registered under the Intra Government heading? I believe the government only holds registered bonds. So I don't think we need to make this distinction.

The different spreadsheets have different parts of this tree. The details can be seen directly in the spreadsheets, in which the type of series is recorded below the Treasury's Name of Issue. Here, we briefly discuss the available information for *Active Outstanding*:

- For the period 1776_1918 there are no *Intra Government-Holdings*. There are data only on the *Active Outstanding* broken down in Coupon and Registered (where appropriate). This data is contained in the spreadsheets beginning with `Gov_Bond_Quant` and the dates inside the range 1776_1918.
- For the period 1919–1960 there are data on *Intra Government-Holdings*. These data are contained in the files beginning with `Gov_Bond_Gov_Accounts`. The data contained in the spreadsheets beginning with `Gov_Bond_Quant` and the dates inside the range 1919_1960 contains the total *Active Outstanding* (no longer broken down into Coupon and Registered).

XXXXX JP: George, are you happy with this description of the data. Please feel free to change anything that you think is not accurate. I am happy with your excellent descriptions. I made a few minor changes.

1.2.2 Series for L1 ID 1 (Pre 1790 Domestic Debt) and L1 ID 4 (Other)

Pre 1790 Domestic Debt and *Other* debt have a much simpler series taxonomy. This is shown in figure 2, where the descriptions of the series types are the same as in section 1.2.1.

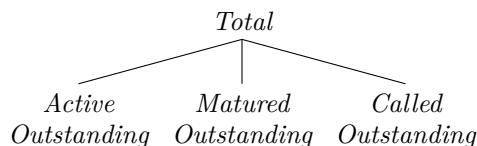


Figure 2: Types of Quantity Series: Pre 1790 Domestic Debt and Other Debt

Pre 1790 Domestic Debt and *Other* debt have only been included in the data set to ensure that the total quantity outstanding adds up correctly. They can be excluded from most analysis.

1.2.3 Series for L1 ID 3 (Non-Interest Bearing)

Non-Interest Bearing debt has two L3 categories: *Commodity Currency* and *Fiat Currency*. *Fiat Currency* has only one series type: *In Circulation*. The series types for *Commodity Currency* are shown in figure 3.

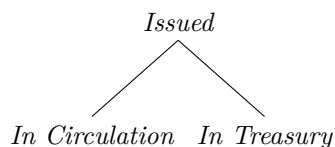


Figure 3: Types of Quantity Series: Non-Interest Bearing Debt

George, do you want to provide a description of these series types or do you think it is self explanatory? This is fine. GJH

1.2.4 Series for L1 ID 5 (Asset)

Bonds in the *Asset* category only have one series: *Total Outstanding*.

1.3 RAWDATA/BONDPRICE FOLDER:

The price time series data for the bonds is spread across the following files:

- Gov_Bond_Prices_1776_1859.xlsx
- Gov_Bond_Prices_1840_1900.xlsx
- Gov_Bond_Prices_1900_1927.xlsx

- Gov_Bond_Prices_1919_1949.xlsx

- Gov_Bond_Prices_1911_1918.xlsx

The prices reported in Gov_Bond_Prices_1776_1859.xlsx are from Raza-ghian [2002] and Sylla, Wilson, and Wright [2006].

The prices reported in Gov_Bond_Prices_1840_1900.xlsx and the Raza-ghian [2002] and from the *Commercial and Financial Chronicle*.

The source of the prices reported in Gov_Bond_Prices_1900_1927.xlsx is the *Commercial and Financial Chronicle*.

The prices reported in Gov_Bond_Prices_1911_1918.xlsx are from the US Treasury Circulars.

The price reported in Gov_Bond_Prices_1919_1949.xlsx are from “United States Govt. Bonds” tables in the New York *Times* until 1925. Thereafter the prices are from the *CRSP US Treasury Database*.²

Each bond has multiple possible types of price series. For each column, the series type is listed below the Treasury’s Name of Issue in the spreadsheets. Figure 4 shows, diagrammatically, how the different possible types of series fit together.

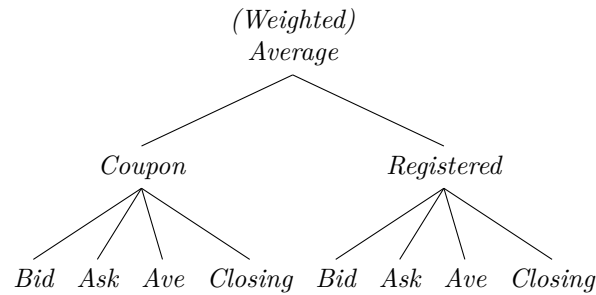


Figure 4: Types of Price Series

The series do not add up directly. Their relations are described below:

- The (Weighted) *Average* price is ‘calculated’ in the following way:
 - If a bond only has one price series, then that series is assigned to the *Average* price series.
 - If the bond has both Coupon and Registered series, then the *Average* series is calculated in python (not in the spreadsheet). It will be calculated using the following rules:
 - * At dates for which there is sufficient information to calculate a quantity weighted average of the Coupon and Registered prices, the quantity weighted average is assigned to the *Average* series.

²See <http://www.crsp.com/products/research-products/crsp-us-treasury-database>.

- * At dates for which there is complete price information on both Coupon and Registered series but incomplete quantity information, an evenly weighted average of the Coupon and Registered prices is assigned to the *Average* series.
- * At dates for which there is exactly one of a Coupon or a Registered price, the existing price is assigned to the *Average* series.
- For each Coupon or Registered price, there are at most four types of available prices: *Bid*, *Ask*, *Ave* and *Closing*. Only one of the four types is selected for each series. Selection has been done with the following order of preference (subject to data availability):
 1. *Closing* price
 2. *Ave* price
 3. *Bid* price
 4. *Ask* price (which typically seems to have the least information)

1.4 RAWDATA/MACRO DATA FOLDER

The spreadsheet `usvar.xlsx` contains times series data for the following macroeconomic time series:

- `gdpNom` : contains nominal GDP (in millions of dollars)
- `gdpReal` : contains real GDP (in millions of 2009 dollars)
- `gdpDeflator` : contains GDP deflator
- `pop` : contains population (in thousands)
- `gdpNomGrowth` : contains annual growth of nominal GDP
- `gdpRealGrowth` : contain annual growth of real GDP
- `inflation` : contains the annual growth of the GDP Deflator
- `popGrowth` : contains the annual growth in the population

The spreadsheet `goldprice.xlsx` contains the time series data for the following series:

- `Price of Gold` : Price of gold in greenbacks from 1862 to 1879. Source: Table 2 of Mitchell [1908]
- `$/Sterling 60 Days` : **XXX**
- `$/Sterling 3 Days` : **XXX**
- `Index of Economic Activity` : **XXX**

- Total Imports : XXX

The source of the data in `usvar.xlsx` is <https://www.measuringworth.com/>. Since we are simply taking data from a readily available source without any value added, I propose cutting these data from our database.

Since we use the price of gold I recommend keeping that series. We do not use the other four series, and I will have to do some digging to confirm the sources. I propose cutting these series. GJH

XXXXX George to fill in missing descriptions on gold prices and briefly describe the different sources that the macro data comes from.

1.5 RAWDATA/ISSUEREDEMPTIONS FOLDER

XXXXX George to add any comments, if required.

2 CODE FOR IMPORTING THE EXCEL FILES

The code for importing the underlying excel files into python pandas is located in the `/Industrial` Folder. The files `RunFile.py` runs all the other import files and exports the final data frames into the folder `/Industrial/DataFrames`.

2.1 NOTES ON IMPORTING THE LIST DATA

XXXXX JP to add as required

2.2 NOTES ON IMPORTING THE QUANTITY DATA

XXXXX JP to add as required

2.3 NOTES ON IMPORTING THE PRICE DATA

The following conventions are used when importing the bond price spreadsheets:

- `Gov_Bond_Prices_1776_1859.xlsx`:
 - Price series are classified using the series type description near the top of each column. (For this spreadsheet, all prices series are set to *Average*.)
 - Only price data from the period 1776-1840 is taken from this spreadsheet.
- `Gov_Bond_Prices_1840_1900.xlsx`:
 - Price series are classified using the series type description near the top of each column.
 - The series with type *Matured* and *Called* are ignored since they are always empty.

- `Gov_Bond_Prices_1900_1927.xlsx`:
 - For all bonds, only the *closing* series is incorporated into the final data frames. The classification as *Coupon* or *Registered* is then made using the description in each column.
- `Gov_Bond_Prices_1911_1918.xlsx`:
 - For all bonds, only the *Ave* series is incorporated into the final data frames. The classification as *Coupon* or *Registered* is then made using the description in each column.
 - The bonds in the `Gov_Bond_Prices_1911_1918.xlsx` spreadsheet are a subset of the bonds in the `Gov_Bond_Prices_1900_1927.xlsx` spreadsheet. When data overlap occurs, the entries in the `Gov_Bond_Prices_1911_1918.xlsx` are used. That is, `Gov_Bond_Prices_1911_1918.xlsx` is dominant when the spreadsheets are merged.
- `Gov_Bond_Prices_1919_1949.xlsx`:
 - The *Avg* price is selected for tabs with no negative *Ask* price information. The *Bid* price is selected for tabs with negative *Ask* price information. This means that:
 - * *Avg* prices are selected for the following tabs: ‘Consols Loans’, ‘TNotes’ and ‘Cert of Indebt’
 - * *Bid* prices are selected for the following tabs: ‘PCanal’, ‘Liberty Loan’, ‘TBonds’ and ‘TBills’
 - Price series are then classified using the series type description near the top of each column. (For this spreadsheet, all prices series are set to *Average*.)
- As discussed, the (Weighted) *Average* price is ‘calculated’ in the following way:
 - If a bond only has one price series, then that series is assigned to the *Average* price series.
 - If the bond has both *Coupon* and *Registered* series, then the *Average* series is calculated using the following rules:
 - * At dates for which there is sufficient information to calculate a quantity weighted average of the *Coupon* and *Registered* prices, the quantity weighted average is assigned to the *Average* series.
 - * At dates for which there is complete price information on both *Coupon* and *Registered* series but incomplete quantity information, an evenly weighted average of the *Coupon* and *Registered* prices is assigned to the *Average* series.
 - * At dates for which there is exactly one of a *Coupon* or a *Registered* price, the price is assigned to the *Average* series.

3 DATA FRAMES

This section describes the location and organization of the final DataFrames produced by the import code. These data frames are labeled:

- **BondList**: which contains list information about each bonds.
- **BondQuant**: which contains times series data on the quantity outstanding for each bond
- **BondPrice**: which contains times series data on the quantity outstanding for each bond
- **MacroData**: which contains time series data on the relevant macroeconomic variables

3.1 STORAGE FORMATS

All the data frames have been saved in three different file formats: `.csv`, `.pkl` and `.hdf5`. However, it should be noted that the versions of the data frames saved in the `.hdf5` format have been modified to exclude any non-numerical or date-time objects. The jupyter notebook `Import_Instructions.ipynb` illustrates how the import the different data frames.

3.2 BONDLIST DATA FRAME

The full **BondList** data frame contains the following columns (with the data type of the column noted in brackets). Unless otherwise stated in the table below, the columns have the same interpretation as was described in section 1.1.

- *Import ID* (int): The unique identifier used to name the bonds in the underlying spreadsheets.
- *Category L1* (str): Same as section 1.1
- *Category L2* (str): Same as section 1.1
- *Category L3* (str): Same as section 1.1
- *Category L1 ID* (int): Same as section 1.1
- *Category L2 ID* (int): Same as section 1.1
- *Category L3 ID* (int): Same as section 1.1
- *Issue L1 ID* (int): Bond issue number when category L1 is used to index the bonds
- *Issue L2 ID* (int): Bond issue number when category L2 is used to index the bonds

- *Issue L3 ID* (int): Bond issue number when category L3 is used to index the bonds
- *L1 ID* (int): Unique bond ID when category L1 is used to index the bonds
- *L2 ID* (int): Unique bond ID when category L2 is used to index the bonds
- *L3 ID* (int): Unique bond ID when category L3 is used to index the bonds
- *Treasury's Name Of Issue* (str): Same as section 1.1
- *Authorizing Act* (str): String describing the authorizing act (not necessarily a date)
- *Term Of Loan* (str): String describing the duration of the loan (not necessarily a number)
- *Issue Date* (pd.Timestamp): Same as section 1.1
- *Authorizing Act Date* (pd.Timestamp): Same as section 1.1
- *First Issue Date* (pd.Timestamp): Same as section 1.1
- *First Redemption Date* (pd.Timestamp): Same as section 1.1
- *Final Redemption Date* (pd.Timestamp): Same as section 1.1
- *Redeemable After Date* (pd.Timestamp): Same as section 1.1
- *Payable Date* (pd.Timestamp): Same as section 1.1
- *Coupon Rate* (float): Same as section 1.1
- *Coupon Rate : Unique* (float): Same as section 1.1
- *Coupon Rate : Min* (float): Same as section 1.1
- *Coupon Rate : Max* (float): Same as section 1.1
- *Coupon Frequency* (str): Same as section 1.1
- *Coupons Per Year* (float): Same as section 1.1
- *CF01* (float): Same as section 1.1
- *CF02* (float): Same as section 1.1
- *CF03* (float): Same as section 1.1
- *CF04* (float): Same as section 1.1
- *CF05* (float): Same as section 1.1
- *CF06* (float): Same as section 1.1

- *CF07* (float): Same as section 1.1
- *CF08* (float): Same as section 1.1
- *CF09* (float): Same as section 1.1
- *CF10* (float): Same as section 1.1
- *CF11* (float): Same as section 1.1
- *CF12* (float): Same as section 1.1
- *Ambiguous* (float): Same as section 1.1
- *Coin* (float): Same as section 1.1
- *Callable* (float): Same as section 1.1
- *Price Sold* (float): Same as section 1.1
- *Authorized Amount* (float): Same as section 1.1
- *Amount Issued* (float): Same as section 1.1
- *Limit On Issued* (float): Same as section 1.1
- *Marginal Increase In Total Debt Authorized* (float): Same as section 1.1

The BondList data frame saved in the `.hdf5` format excludes any columns with strings.

In all formats, the BondList data frame has been saved with the column *L3 ID* as the index.

3.3 BONDQUANT DATA FRAME

The BondQuant data frame contains all the available quantity series for each bond. The BondQuant columns have two layers or multi-indexing. The first layer is the identification number of the bond (*L3 ID* from the BondList data frame). The second layer is the series type. `BondQuant.csv` and `BondQuant.pkl` record the series type using a string description. The `.hdf5` file record the series type using the following numerical scheme:

1. Total Outstanding
2. Active Outstanding
3. Matured Outstanding
4. Called Outstanding
5. Public Holdings
6. Intra Government Holdings

7. Coupon
8. Registered
9. Issued
10. In Circulation
11. In Treasury

3.4 BONDPRICE DATA FRAME

The BondPrice data frame contains all the available price series for each bond. The BondPrice columns have two layers or multi-indexing. The first layer is the identification number of the bond (*L3 ID* from the BondList data frame). The second layer is the series type. `BondQuant.csv` and `BondQuant.pkl` record the series type using a string description. The `.hdf5` file record the series type using the following numerical scheme:

1. Average
2. Coupon
3. Registered

References

- Rafael A. Bayley. *National Loans of the United States*. Government Printing Office, Washington, 1882.
- Board of Governors of the Federal Reserve System. *Banking and Monetary Statistics, 1914-1941*. Board of Governors of the Federal Reserve System, Washington D.C., 1943.
- C.F. Childs. *Concerning U.S. Government Securities: A Condensed Review of the Nation's Currency, Public Debt, and the Market for Representative United States Government Loans, 1635-1945. Also a Chronology of Government Bond Dealers*. The Lakeside Press, R.R. Donnelley & Sons Company,, Chicago, IL, 1947.
- W.C. Mitchell. *Gold, Prices, and Wages Under the Greenback Standard*. Berkeley. University of California. Publ. in economics. University Press, 1908.
- Rose Razaghian. Financial credibility in the united states: The impact of institutions, 1789-1860. Data made available through the International Center for Finance at Yale University., 2002.
- U.S. Treasury Register's Office. *Statement of Receipts and Expenditures of the Government (by Warrants) from July 1, 1855, to June 30, 1885 and Statement of Principal of Public Debt from 1791 to 1836 by Issues and Redemptions, and From January 1, 1836, to June 30, 1885 by Warrants*. Government Printing Office, Washington DC, 1886.
- Richard E. Sylla, Jack Wilson, and Robert E. Wright. Early u.s. security prices. <http://eh.net/databases/early-us-securities-prices>, 2006.
- United States Department of the Treasury. *Monthly Statement of the Public Debt*. U.S. Treasury Department, 1869-2015. URL <https://www.treasurydirect.gov/govt/reports/pd/mspd/mspd.htm>.
- United States Department of the Treasury. *Treasury Bulletin*. U.S. Treasury Department, Office of the Secretary., 1939-2015. URL <https://fraser.stlouisfed.org/title/407>.