A Data Construction

A.1 Income Share of Finance

Figure 1 shows the income share of the financial sector using various methods and sources to arrive at four series. The two older series, "VA fire, Hist" and "WN fire, hist" categorize the financial sector as including finance, insurance and real estate. The underlying data for these series come from the Historical Statistics of the United States[1] (HSUS). The two more recent series, "WN fin, NIPA" and "VA fin, NIPA", exclude real estate from their definition of the financial sector, and the source for these series is the Bureau of Economic Analysis[2]. These series are frequently revised, and the data used in the paper was current as of May 11, 2011.

The size of the financial sector is measured both by the ratio of its value added relative to GDP (in the series labeled with VA) and by the ratio of total employee compensation in the financial sector relative to aggregate compensation (in the series labeled WN).

Because these series all appear to follow each other closely over the dates in which they overlap, they can be merged into one composite measure representing the GDP share of the financial sector for the full time series. This composite series is displayed in the left hand panel of Figure 2. It is constructed by using the modern, value-added series from 1955-2009. From 1929 to 1955 the GDP share series is estimated using the modern employee compensation data.

Splicing in the earlier data requires an assumption regarding the relative size of the contribution of real estate, since that is included within the definition of the financial sector in the older data taken from the HSUS. This adjustment is made using the detailed industry data available from Kuznets[3] covering 1919-1938. In our analysis, the historical wage series ("WN fire, hist") is appended to our composite series over the period 1900-1928, the adjustments suggested from comparison with Kuznets' series scales down the series by 12.77%. Prior to 1900, the "VA fire, Hist." measure is used, scaled down by 12.32%.

While this GDP share variable looks at the size of the financial sector relative to the entire domestic product, the other data series shown in the left panel of Figure 2 consider alternative denominators that exclude defense spending or that exclude defense and farming. The right hand side shows two additional alternatives. The data series labeled "Domestic" removes net exports of financial services from the numerator, leaving aggregate GDP in the denominator. The data series labeled "Share of Services" takes the previously calculated numerator (without export adjustments) and considers only the value added for services industries in the denominator. Note that the calculation of the GDP share for services follows the same methodology of using employee compensation for older dates where value added measures are not available.

A.2 Credit Markets

The paper shows measures for debt levels for farms, households and non-farm business. These are displayed in Figure 3. Additionally, flow variables are constructed measuring gross corporate bond issuance and the flow of debt to households.

The debt level series come from the Federal Reserve's Flow of Funds data from 1945 to 2010, scaled by the previously discussed GDP series. Historical data on public and private debt from 1916 to 1945 comes from the HSUS, which derives household and farm debt discontinued series published in the "Survey of Current Business" by the Office of Business Economics (which is now the U.S. Bureau of Economic Analysis). The calculation of non-farm business debt also comes through the HSUS, where long-term and short-term debt are extracted from corporate tax return data from 1926 to 1945. These debt levels, divided by GDP, are plotted in Figure 3.

Two flow variables are shown in Figure 4. The "Gross Corp. Bond Issuance" variable scales bond issuance of non-financial corporations by GDP. The primary source is Baker and Wurgler (2000)[4], who provide data

[1] Carter, Susan B. 2006. Historical statistics of the United States : earliest times to the present. Cambridge ;New York : Cambridge University Press. Digital version of dataset available at: http://hsus.cambridge.org

[2] Available online at http://www.bea.gov/industry/gdpbyind_data.htm.

[3] Kuznets, S. (1941): "National Income and Its Composition, 1919-1938," Discussion paper, National Bureau of Economic Research.

[4] Data available online from Wurgler's website at: http://www.stern.nyu.edu/~jwurgler/data/Equity_Share_3.xls

from 1927 to 2008. Historical data from HSUS is appended from 1910-1926,[5] and more recent data from the Securities Industry and Financial Markets Association (SIFMA) covers 2009 and 2010. For portions of the historical data in which total bond issuance includes financial firms, the series is scaled down by 25%, which is the relatively stable fraction of financial bonds relative to total issuance prior to 1980. More recently, the data suggests financial firms issue approximately half of all bonds.

A.3 Equity Markets

The paper discusses three variables measuring the size and growth of the equity market, two of which are graphed in Figure 5. The first measure looks at total market value relative to GDP, where the total market value of non-financial firms comes from the Flow of Funds data from 1946 to 2010. This can be supplemented with equity data from CRSP from the period 1927 to 1945. Since the total market value of equity reported in the Flow of Funds data is about 38% higher than what is reported in CRSP over their overlapping period from 1946 to 1960, the historical CRSP values are scaled up correspondingly. The series can be extended back to 1886 using data compiled by Jovanovic and Rousseau (2005), who aggregate market value from firm-level data reported in contemporary publications.

The second measure looks at gross equity issuance of non-financial firms. Similar to the debt issuance data, the primary source for the series is Baker and Wurgler (2000) and the historical data (prior to 1960) that does not explicitly exclude the equity issuance of financial firms is scaled down by 25%. Historical data from 1909 to 1926 comes from HSUS. Also mirroring the data construction for debt issuance, SIFMA data is used to cover more recent equity issuance (after 2007). The SIFMA series includes financial firm equity issuance and is scaled down by 60% to match the magnitude of non-financials in their overlapping period. The composite series is divided by GDP and smoothed with using a 3-year moving average. This series is plotted in Figure 5 along with the market value of equity over GDP.

The third measure focuses specifically on IPOs. The IPO issuance data is a composite of data from SIFMA for the year 2011, data from Ritter (2011) for 1990-2010, and an adjusted version of the data collected by Jovanovic and Rousseau (2005) for the period 1886 to 1989. Similar to the equity issuance variable, IPO issuanceis divided by GDP and smoothed using a 3-year moving average and plotted alongside the time series of the market value of equity over GDP, as seen in Figure 10.

A.4 Deposits

The deposits of households, firms and local governments are collected from the Flow of Funds data, which is available from 1946 to 2010. As described in the text, these deposits include all currency, deposits, money market funds, and somewhat less traditionally, non-financial repos. Local government deposits recorded in the Flow of Funds data in the year 1945 appear anomalously large, so this data point is excluded. Data from 1867 to 1945 is derived from the HSUS data. The quantity of corporate and government deposits is estimated by subtracting household deposits from the M3 definition of money supply. Over the time period where the HSUS data overlaps the Flow of Funds data, this methodology seems to be approximately 15% larger than the deposits figure from the Flow of Funds data, so the historical series is scaled down by a factor of 1.15. These various series are plotted in Figure 6.

A.5 Financial Output Measures

The paper creates two aggregated measures of financial output, a series based on output and a series based on flows. These are plotted in the left panel of Figure 7. Each of these are an aggregation of six components,

$$y^{\phi} = b_c + d_k + e_k + 0.1 \times b_g + y_{M\&A}.$$

For the level series, the component labeled b_c corresponds to household borrowing and uses the household debt series described in the Credit Markets subsection and plotted in Figure 3. The series in the figure only goes back to 1916, but this is extended even further by incorporating historical mortgage data from Schularick and Taylor (forthcoming) and predicting household debt values from 1890 to 1915 by using the relationship

[5] Corporate security issues: 1910-1934, series Cj831-837

from a regression of their overlapping values from 1916 to 1976 (which roughly corresponds to increasing the mortgage debt by 50% and adding a constant nominal term of $52 million). Lastly, an additional 30 years is added to the historical series by taking the time series of aggregated assets on the balance sheets of financial institutions, using predicted relationships to estimate the portion that is private lending, and then estimating that 31% of this lending is to households (which is the exact ratio that would align the data in 1890). The flow series for household debt uses Flow of Funds data when available (1946 to 2010) and otherwise calculates the flow as the growth in the level series (relative to GDP). All flows must be positive, and this is ensured by censoring at 0.1% of GDP. This censoring affects 13% of years in the flow time series. The final flow series is then smoothed using a three-year moving average.

The components labeled d_k and e_k correspond to business debt and equity. The debt level series combines the farm and non-farm business borrowing series described in the Credit Markets subsection above, and the equity level series is described in the Equity Markets subsection. Like household debt, the business debt series is expanded historically back to 1860 by using historical information on the loans carried on the balance sheets of financial firms, assuming 66% of this lending is to businesses and subtracting any foreign debt financing. An additional adjustment is made to adjust for the non-corporate sector, as described in section 3.4 of the paper. The construction of the flow series for business debt and equity uses the same censoring and smoothing procedure described for households.

The final three components are government debt (b_g), deposits (m), and mergers and acquisitions ($y_{M\&A}$). As discussed in the paper, the contribution of government debt, b_g, to financial output is assumed to one tenth the cost of private debt for both flow and issuance. Deposits are only measured as a level series, and the construction of this component was previously described. Merger and acquisition activity, on the other hand, is only measured as a level series, and the construction of this series is described in detail in section 3.4 of the paper.

A composite of the level and flow series is created by averaging the two, where flows are scaled by a factor of 8.48 to make the two series comparable. The components of the composite series are plotted in the right hand panel of Figure 7.

A.6 Intermediation Cost

A time series of the unadjusted intermediation cost of the financial sector is plotted in Figure 8. It is simply the ratio of the domestic income share of the financial sector, plotted on the right-hand panel of Figure 2, over the composite output measure of the financial sector, combining the flow and level output measures plotted in Figure 7 and as described above.

A.7 Quality Adjustments

Section 4 of the paper discusses the possibility that the financial sector may face increasing marginal costs in supplying credit to riskier firms or households. The adjustments are displayed on Figure 11 of the paper. The fraction of low wealth households is based on the time series of the income share of those above the 90th percentile from Piketty and Saez (2003), which will be labeled $saez_{p90}$. The Piketty and Saez time series is available online[6] 1917 to 2008. Prior to 1917 the value is assumed to be 0.385 and after 2008 the value is assumed to be 0.455. The associated quality adjusted output series are plotted in Figure 11. By dividing the cost of financial intermediation by the quality-adjusted output of the financial sector, we get the quality-adjusted cost index displayed in Figure 12.

[6] from Saez's homepage http://elsa.berkeley.edu/~saez/ as of November 2011.