

Has the US Finance Industry Become Less Efficient? On the Theory and Measurement of Financial Intermediation[†]

By THOMAS PHILIPPON*

A quantitative investigation of financial intermediation in the United States over the past 130 years yields the following results: (i) the finance industry's share of gross domestic product (GDP) is high in the 1920s, low in the 1960s, and high again after 1980; (ii) most of these variations can be explained by corresponding changes in the quantity of intermediated assets (equity, household and corporate debt, liquidity); (iii) intermediation has constant returns to scale and an annual cost of 1.5–2 percent of intermediated assets; (iv) secular changes in the characteristics of firms and households are quantitatively important. (JEL D24, E44, G21, G32, N22)

This paper is concerned with the theory and measurement of financial intermediation. The role of the finance industry is to produce, trade, and settle financial contracts that can be used to pool funds, share risks, transfer resources, produce information, and provide incentives. Financial intermediaries are compensated for providing these services. The income received by these intermediaries measures the aggregate cost of financial intermediation. This income is the sum of all spreads and fees paid by nonfinancial agents to financial intermediaries and it is also the sum of all profits and wages in the finance industry. This cost of financial intermediation affects the user cost of external finance for firms who issue debt and equity, and the costs for households who borrow or use asset management services.

* Stern School of Business, New York University, 44 West Fourth Street, New York, NY 10012, NBER, and CEPR (e-mail: tphilipp@stern.nyu.edu). This has been a very long project. The first draft dates back to 2007, with a focus on corporate finance, but without the long-term historical evidence. This paper really owes a lot to other people, academics and non-academics alike. Darrell Duffie, Robert Lucas, Raghuram Rajan, Jose Scheinkman, Robert Shiller, Andrei Shleifer, and Richard Sylla have provided invaluable feedback at various stages of this project. Boyan Jovanovic, Peter Rousseau, Moritz Schularick, and Alan Taylor have shared their data and their insights, and I have greatly benefited from discussions with Lewis Alexander, Patrick Bolton, Markus Brunnermeier, John Cochrane, Kent Daniel, Douglas Diamond, John Geanakoplos, Gary Gorton, Robin Greenwood, Steve Kaplan, Anil Kashyap, Ashley Lester, Andrew Lo, Guido Lorenzoni, Andrew Metrick, William Nordhaus, Matthew Rhodes-Kropf, David Robinson, Kenneth Rogoff, Tano Santos, David Scharfstein, Hyun Shin, Jeremy Stein, Gillian Tett, Wallace Turbeville, and Luigi Zingales, as well as seminar participants at Stanford, Yale, NYU, Harvard, Chicago, Princeton, and the Paris School of Economics. I also thank Paul Krugman for his discussion at the 2011 NY Area Monetary conference, Axelle Ferriere, Peter Gross, Andrea Prestipino, Markus Siivonen, Robert Turley, and Shaojun Zhang for research assistance, and the Smith Richardson Foundation for its financial support. The author declares that he has no relevant or material financial interests that relate to the research described in this paper. Two online appendices describe calibrations and data sources.

[†]Go to <http://dx.doi.org/10.1257/aer.20120578> to visit the article page for additional materials and author disclosure statement.

In equilibrium, the user cost of external finance is the sum of the rate of returns to saver (r) and the unit cost of financial intermediation (ψ):

$$(1) \quad \text{user cost of finance} = r + \psi.$$

The unit cost of intermediation ψ can in turn be measured as the ratio of the income of financial intermediaries to the quantity of intermediated assets. The goal of this paper is to construct these three measures: the income of financial intermediaries, the quantity of intermediated assets, and, finally, the unit cost of intermediation ψ .

There are several motivations for undertaking such a project. A first motivation is that ψ has a direct impact on the overall efficiency of the economy. Even small changes in ψ have large long run effects on the capital/output ratio, and therefore on income per capita. Equations such as (1) play a central role in the literature that seeks to quantify the consequences of financial development for economic growth.¹

A second motivation is to shed light on the transformation of the finance industry that has occurred since the 1970s. For instance, we would certainly like to know if the move away from traditional banking and toward an “originate-and-distribute” model has lowered the cost of funds for households and businesses. This is precisely what ψ should measure. Similarly, if derivatives markets lower hedging costs, their growth should translate into lower funding costs and higher asset values. Any debate about financial regulation is also a debate about ψ , since it involves a trade-off between safety and efficiency.² The broader point here is that learning about ψ is important from a positive perspective and from a normative perspective.

This paper seeks to define and measure financial intermediation. It treats the finance industry as a black box and attempts to measure what goes in, what comes out, and how much the whole system costs. It is important, however, to understand just how difficult the measurement problem is. A simple illustration is given in Figure 1. At the prevailing market rates of 5 and 7 percent, borrowers (firms or households) want to borrow \$100, and savers want to save \$100. To flow back and forth between savers and borrowers, the funds go through financial intermediaries. These intermediaries need \$2 to pay their wage bill and rent the necessary capital. In the terminology of this paper, the quantity of intermediated assets is \$100, the intermediation cost is \$2, and therefore $\psi = 2$ percent. Figure 1 presents two fundamentally equivalent ways to organize financial intermediation. In traditional banking, intermediation occurs under one roof: the bank makes a loan, keeps it on its books, and earns a net interest income. This income compensates for the cost of screening and monitoring the borrower and for managing the duration and credit

¹ See Greenwood, Sanchez, and Wang (2010); Buera, Kaboski, and Shin (2011); and Midrigan and Xu (2014) for recent analyses of financial development and growth. In addition, much of recent work has focused on the macroeconomic consequences of a sudden increase in ψ , and on the link between ψ and intermediary capital, leverage, and liquidity. Cúrdia and Woodford (2010); Gertler and Kiyotaki (2010); Hall (2011); Christiano and Ikeda (2011); and Corsetti et al. (2011) study the impact of negative shocks to financial intermediation, building on the classic contribution of Bernanke, Gertler, and Gilchrist (1999). Gertler and Karadi (2011); He and Krishnamurthy (2012); and Moore (2011) focus on liquidity. This paper only deals with the long-term evolution of ψ , but the value of ψ in normal times is an important parameter even if one is interested in the deviations from its long-term trend.

² Take the debate about capital adequacy ratios for banks for instance. Bank leverage has risen substantially since the late nineteenth century, as discussed in Haldane, Brennan, and Madouros (2010) among others. If Modigliani and Miller’s (1958) proposition holds, as Admati et al. (2011) argue, we should not expect a link between ψ and bank leverage, but if the proposition fails we might expect a downward drift in ψ over time.

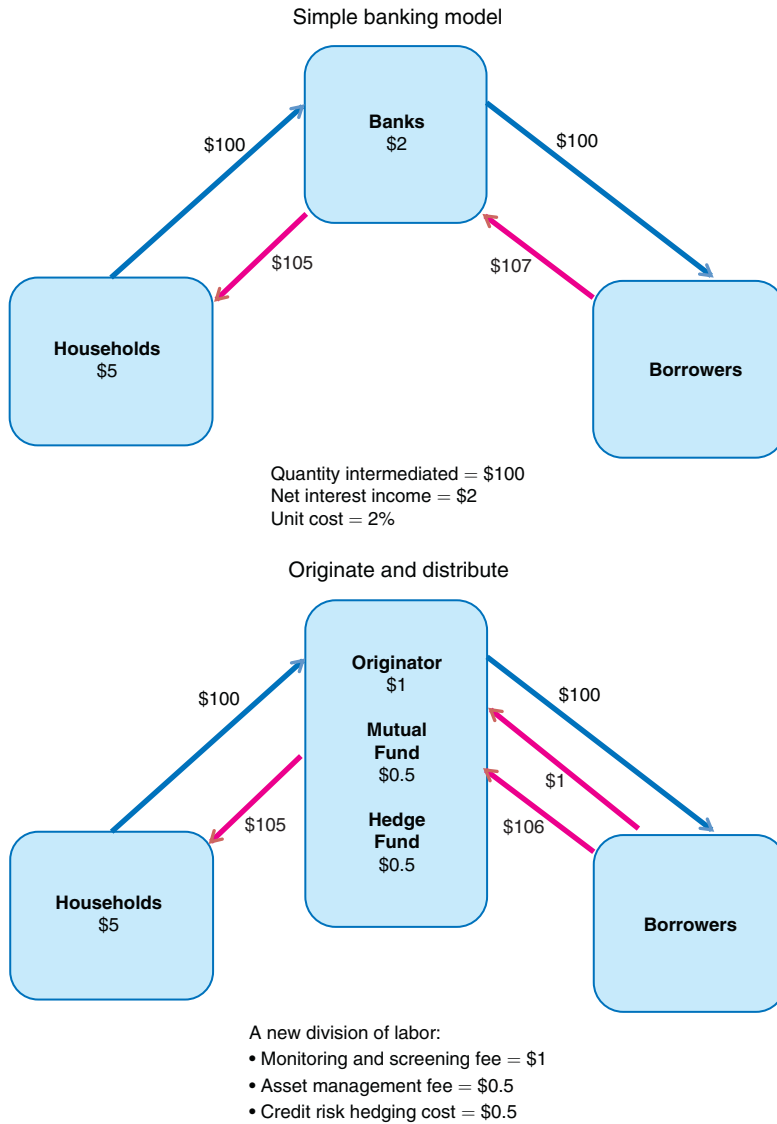


FIGURE 1. TWO EQUIVALENT MODELS OF FINANCIAL INTERMEDIATION

risk of the loan. In the originate and distribute model, by contrast, there is a daisy chain of intermediation. Many transactions occur *inside* the black box, with total face values potentially much larger than \$100. There is no simple measure of net interest income as in the traditional model: there are origination fees, asset management fees, trading profits, etc. But the sum of wages and profits for all intermediaries is still \$2, and the quantity of intermediated assets seen from *outside* the black box is still \$100.

Three steps are required, then, to understand financial intermediation, and these steps determine the structure of the paper: (i) measure the income of financial intermediaries; (ii) define and construct the quantity of intermediated assets; and (iii) compute the unit cost of intermediation and perform quality adjustments. For

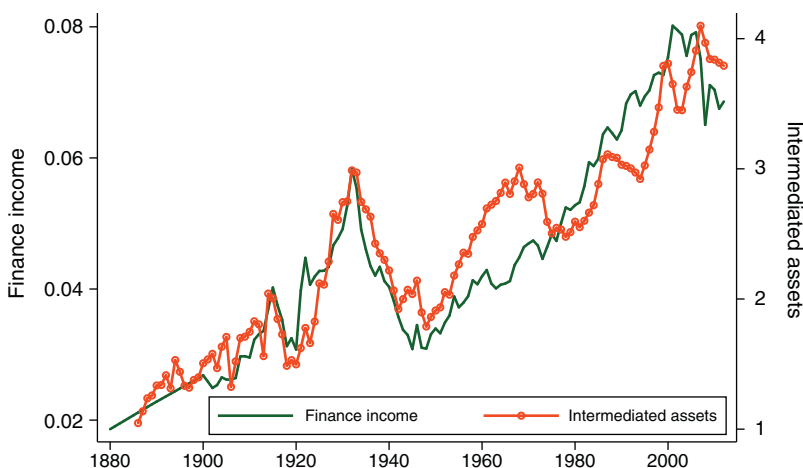


FIGURE 2. FINANCE INCOME AND INTERMEDIATED ASSETS OVER GDP

Notes: Both series are expressed as a share of GDP, excluding defense spending. Finance income is the domestic income of the finance and insurance industries, i.e., aggregate income minus net exports. It is available from 1880 to 2012. Intermediated assets include debt and equity issued by nonfinancial firms, household debt, and various assets providing liquidity services. Data range for intermediated assets is 1886–2012.

the purposes of this paper, the real difficulty lies in the heterogeneity among intermediated assets. Savings vehicles are heterogeneous: perhaps households save \$50 in liquid claims with a return of 4 percent and \$50 in illiquid claims with a return of 6 percent. The average (expected and risk adjusted) return is 5 percent, but intermediaries must bear the cost of creating and maintaining liquid claims. Borrowers are also heterogeneous: young firms and blue chip companies, wealthy households and poor households. Changes in the composition of borrowers affect the cost of intermediation, while improvements in financial intermediation give access to credit to borrowers who were previously priced out.

The first contribution of this paper is empirical. Figure 2 shows that the quantity of financial intermediation varies dramatically over time.³ The first series, constructed in Section I, is the income of financial intermediaries divided by GDP. The income share grows from 2 to 6 percent from 1880 to 1930. It shrinks to less than 4 percent in 1950, grows slowly to 5 percent in 1980, and then increases rapidly after 1980.

Given these large historical variations in the finance income share, it is natural to ask if there are commensurate changes in the quantity of intermediated assets.⁴ Section II relies on a simple extension of the neoclassical growth model

³Note that I use the recently updated data from the Bureau of Economic Analysis. The 2014 Comprehensive Revision of Industry Accounts has led to downward revisions in the estimated value added of finance and insurance for the 1997–2012 period. See <http://www.bea.gov/industry/index.htm>. This revision has occurred between the first version of this paper and the current version, so readers of earlier versions might notice some differences.

⁴We can think of the finance industry as providing three types of services: (i) liquidity (means of payments, cash management); (ii) transfer of funds (pooling funds from savers, screening, and monitoring borrowers); (iii) information (price signals, advising on M&As). Financial firms typically produce a bundle of such services. For instance, risk management uses all three types of activities. Services of type (i) and (ii) typically involve the creation of various financial assets and liabilities. This classification is motivated by the mapping between theory

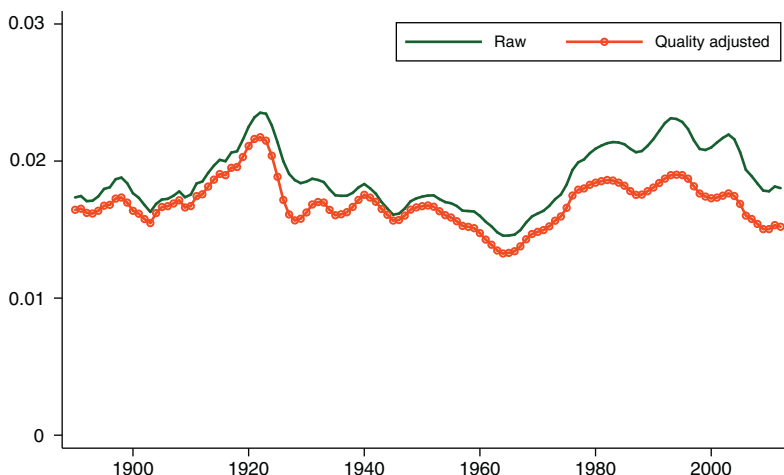


FIGURE 3. UNIT COST OF FINANCIAL INTERMEDIATION

Notes: The raw measure is the ratio of finance income to intermediated assets displayed in Figure 1. The quality adjusted measure takes into account changes in firms' and households' characteristics. Data range is 1886–2012.

as an accounting framework for household finance, corporate finance, and liquidity provision. The size of the various markets varies significantly over time. The most important trend in credit markets in recent years is the increase in household debt. The business credit market is relatively large in the 1920s, small in the 1960s, and large again after 1980, although not as large as in the late 1920s. I also measure the market value of outstanding equity and the flows of initial and seasoned offerings. Deposits, repurchase agreements, and money markets mutual funds are used to measure liquidity services. After aggregating the various types of credit, equity issuances, and liquid assets into one measure, I obtain the quantity of financial assets intermediated by the financial sector displayed in Figure 2.

I can then divide the income of the finance industry by the quantity of intermediated assets to obtain a measure of the unit cost ψ . Figure 3 shows that this unit cost is around 1.5–2 percent and relatively stable over time. In other words, I estimate that it costs \$0.02 per year to create and maintain \$1 of intermediated financial asset. I also find clear evidence that financial services are produced under constant returns to scale. For instance, from 1947 to 1973 (a period of stable growth without major financial crises), real income per capita increases by 80 and real financial assets by 250 percent, but my estimate of the unit cost of intermediation remains remarkably constant.

The raw measure of Figure 3, however, does not take into account changes in the characteristics of borrowers. The final contribution of the paper is to perform quality adjustments to the quantity of intermediated assets. The 1920s and 1990s

and measurement discussed throughout the paper. It differs a little bit from that of Merton (1995). I do not attempt in this paper to measure the informativeness of prices. This issue is tackled by Bai, Philippon, and Savov (2011). See the discussion at the end of Section III.

are times of entry by young and risky firms, and Jovanovic and Rousseau (2005) have shown that this pattern is related to waves of technological innovation. In the household credit market, relatively poor households have gained access to credit in recent years. In both cases, the challenge is to account for the fact that these borrowers require more intermediation per unit of credit extended. I rely on theory to make the required quality adjustments, which appear to be quantitatively important. According to my calculations, in the 1990s, the raw measure of intermediation underestimates the true quantity by about 25 percent. Given the size of intermediated markets, the failure to adjust for quality would represent a measurement error of the order of one GDP. Figure 3 shows that the adjusted unit cost is more stable than the unadjusted one.

Even with the quality adjustment, however, I find that the unit cost of intermediation is about as high today as it was at the turn of the twentieth century. Improvements in information technologies do not appear to have led to a significant decrease in the unit cost of intermediation. Explaining this puzzle is an active area of research, some of which is discussed at the end of Section III.

Related Literature.—Financial intermediation does not have a benchmark quantitative model in the way asset pricing does. By using a model to interpret long time series of prices and quantities, and by providing a set of stylized facts for future research, this paper shares the spirit of Mehra and Prescott (1985). But because financial intermediation is a more heterogeneous field than asset pricing, this paper has to draw from several strands of the literature in finance and economics.

The first strand is the theory of banking and financial intermediation. While stylized and focused on macroeconomic predictions, the model developed below is consistent with leading theories of financial intermediation, such as Diamond and Dybvig (1983); Diamond (1984); Gorton and Pennacchi (1990); Holmström and Tirole (1997); Diamond and Rajan (2001); and Kashyap, Rajan, and Stein (2002). Gorton and Winton (2003) provide a review of the literature on financial intermediation. However the focus of this paper differs from that of the intermediation literature in several ways. First, I focus on the measurement of intermediation costs. Second, I model household and corporate finance simultaneously. Third, I use an equilibrium model to give a quantitative interpretation of the historical evidence.

There is a large literature on financial development, which I do not have room to discuss here, except to say that it tends to focus on cross-sectional comparisons of countries at relatively early stages of financial development in order to understand the impact of finance on economic growth (e.g., Rajan and Zingales 1998), and the determinants of financial development itself (e.g., La Porta et al. 1998; Guiso, Sapienza, and Zingales 2004). The literature typically focuses on corporate finance (Greenwood, Sanchez, and Wang 2010; Buera, Kaboski, and Shin 2011; Midrigan and Xu 2014); except Mehra, Piguillem, and Prescott (2011), who study intermediation in a model where households save for retirement over an uncertain lifetime.⁵

⁵My approach is complementary to this literature and uses many of its important insights. The difference is that I focus on the evolution of the entire US finance industry. As a result, both theory and measurement must be expanded. For instance, following Beck, Demirgüç-Kunt, and Levine (2001), the literature uses cross-country data on interest-rate spreads to estimate financing frictions (e.g., Greenwood, Sanchez, and Wang 2013). To study the US finance industry, it is important to recognize that non-interest income (fees, trading revenues, etc.) is now the

This paper is more closely related to a recent branch of the literature that seeks to provide risk-adjusted measures of financial productivity (Wang, Basu, and Fernald 2009; Haldane, Brennan, and Madouros 2010; Basu, Inklaar, and Wang 2011).

In its account of liquidity services provided by the finance industry, this paper is also related to the classic literature on money and banking. Lucas (2000) provides an analysis of money demand. Kiyotaki and Moore (2008) study the interaction of liquidity, asset prices, and aggregate activity. A recent branch of this literature has focused on the rise of market-based intermediation, also called shadow banking. Pozsar et al. (2010) describe the structure of shadow banking. Gorton and Metrick (2012); Stein (2012); Gorton, Lewellen, and Metrick (2012); and Gennaioli, Shleifer, and Vishny (2013) emphasize the importance of investors' demand for safe assets as a driver of shadow banking activity.

Finally, there is an emerging literature on the growth of the finance industry.⁶ Philippon and Reshef (2012) share the historical perspective of this paper but focus on the composition of the finance labor force. Greenwood and Scharfstein (2013) provide an illuminating study of the growth of modern finance in the United States. They show that two activities account for most of this growth over the past 30 years: asset management and the provision of household credit. For asset management, they uncover an important stylized fact: individual fees have typically declined but the allocation of assets has shifted toward high fee managers in such a way that the average fee per dollar of assets under management has remained roughly constant. While most of the existing work has focused on the United States, Philippon and Reshef (2013) and Bazot (2013) provide evidence for other countries.

A second set of papers offers theoretical explanations for the growth of finance documented in this paper and in the empirical papers discussed above. There are two main stylized facts to explain: the size of finance (see Figure 2) and the unit cost (see Figure 3). Regarding this second stylized fact, a puzzle seems to be that the unit cost has not declined despite obvious improvements in information technologies. As a result, the income received by financial intermediaries might be unexpectedly high. In Glode, Green, and Lowery (2012), an "arms race" can occur as agents try to protect themselves from opportunistic behavior by (over-)investing in financial expertise. In Bolton, Santos, and Scheinkman (2011), cream skimming in one market lowers assets quality in the other market and allows financial firms to extract excessive rents. In Pagnotta and Philippon (2011) there can be excessive investment in trading speed because speed allows trading venues to differentiate and charge higher prices. Gennaioli, Shleifer, and Vishny (2014b) propose an alternative interpretation for the relatively high cost of financial intermediation. In their model, trusted intermediaries increase the risk tolerance of investors, allowing them to earn higher returns. Because trust is a scarce resource, improvements in information technology do not necessarily lead to a lower unit cost.

dominant source of income for financial firms (even for banks: see JPMorgan's 2010 annual report for instance), that consumer credit is at least as important as corporate credit, and that the shadow banks' creation of safe assets is driven by investors' liquidity demand (all these points are discussed in details below).

⁶The large historical changes in the finance share of GDP were first documented and discussed in Philippon (2008), but that paper only focused on corporate credit. The paper did not consider household credit, and did not account for liquidity services, which have become important with the rise of the shadow banking system.

The other fact to explain is the size of the finance industry. Since the unit cost appears to be roughly constant, the question becomes: how do we explain the large historical variations in the ratio of intermediated assets over GDP? This paper documents that the income share of the finance industry is roughly equal to 2 percent of the ratio of intermediated assets over GDP, but it does not seek to explain the size of intermediated assets.⁷ Gennaioli, Shleifer, and Vishny (2014a) propose an explanation. They argue that the growth of finance can be explained by the rise of the wealth to income ratio, documented in Piketty and Zucman (2014) for several countries. The driving force is a slowdown in aggregate growth which leads, along the transition path, to an increase in the capital output ratio. If the unit cost of intermediation does not fall as the capital output ratio increases, then the income share of the finance industry increases.

Let me end this introduction with an important caveat: this paper does not analyze financial crises. The model assumes that credit markets clear via prices, not via covenants or quantity restrictions as we often see during crises. In the model, borrowers can be priced out, but inefficient rationing does not occur. Similarly, the model does not study whether borrowing is appropriate or excessive, whether financial intermediaries take on too much aggregate risk, nor whether government interventions create moral hazard.⁸

The remainder of the paper is organized as follows. Section I estimates the income of financial intermediaries. Section II computes the quantity of intermediated assets. Section III implements quality adjustments and discusses the role of information technology and price informativeness. Section IV concludes.

I. Income Share of Finance

In this section, I present the first main empirical fact: the evolution of the total cost of financial intermediation in the United States over the past 140 years. As argued in the introduction, there is no simple way to break down the income earned by the finance industry into economically meaningful components. For instance, insurance companies and pension funds perform credit analysis, fixed income trading provides liquidity to credit markets, and securitization severs the links between assets held and assets originated. From a historical perspective, these issues are compounded by regulatory changes in the range of activities that certain intermediaries can provide. Rather than imposing arbitrary interpretations on the data, I therefore focus on a consolidated measure of income earned by all financial intermediaries, irrespective of whether they are classified as private equity funds, commercial banks, insurance companies, or anything else.

⁷The household credit model of Section III can “account” for some (but not all) of the rise in consumer debt due to improvements in access to credit. But even there, the goal is not to explain the size of the market, but rather to refine the measurement of the unit cost by removing the bias created by time-varying fixed costs.

⁸See for instance Adrian and Shin (2014); Reinhart and Rogoff (2009); Krishnamurthy (2010); Acharya et al. (2009); and Scharfstein and Sunderam (2011) for recent discussions of these issues.

A. Raw Data

The paper uses a lot of data sources. To save space, all of the details regarding the construction of the series are provided in a separate online Appendix. I focus on the following measure:

$$\frac{y_t^f}{y_t} = \frac{\text{Value Added of Finance Industry}}{\text{GDP}}.$$

Conceptually, the best measure is value added, which is the sum of profits and wages. Whenever possible, I therefore use the GDP share of the finance industry, i.e., the nominal value added of the finance industry divided by the nominal GDP of the US economy. One issue, however, is that before 1945 profits are not always properly measured and value added is not available. As an alternative measure I then use the labor compensation share of the finance industry, i.e., the compensation of all employees of the finance industry divided by the compensation of all employees in the US economy.

Figure 4 displays various measures of the share of the finance and insurance industry in the GDP of the United States estimated from 1870 to 2012. For the period 1947–2012, I use value added and compensation measures from the Annual Industry Accounts of the United States, published by the Bureau of Economic Analysis (BEA). For the post-war period, the two measures display the same trends. This means that, in the long run, the labor share in the finance industry is roughly the same as the labor share in the rest of the economy (in the short run, of course, profit rates can vary). For 1929–1947, I use the share of employee compensation because value-added measures are either unavailable or unreliable. For 1870–1929, I use the Historical Statistics of the United States (Carter et al. 2006).⁹

There are three important points to take away from Figure 4. First, the finance income share varies a lot over time. Second, the measures are qualitatively and quantitatively consistent. It is thus possible to create one long series simply by appending the older data to the newer ones. Third, finance as a share of GDP was smaller in 1980 than in 1925. Given the outstanding real growth over this period, it means that finance size is not simply driven by income per capita.

B. Adjusted Measures

Before discussing theoretical interpretations it is useful to present adjusted series that take into account wars, globalization, and the rise in services.

Wars.—During peace time and without structural change, it would make sense to simply use GDP as the relevant measure of total income. Two factors can complicate the analysis, however. First, WWI and WWII take resources away from the normal production of goods and services. Financial intermediation should then be compared to the non-war-related GDP. To do so, I construct a measure of GDP excluding

⁹Other measures based on Martin (1939) and Kuznets (1941) give similar values. More details regarding the various data sources can be found in Philippon and Reshef (2012) and in the online Appendix.

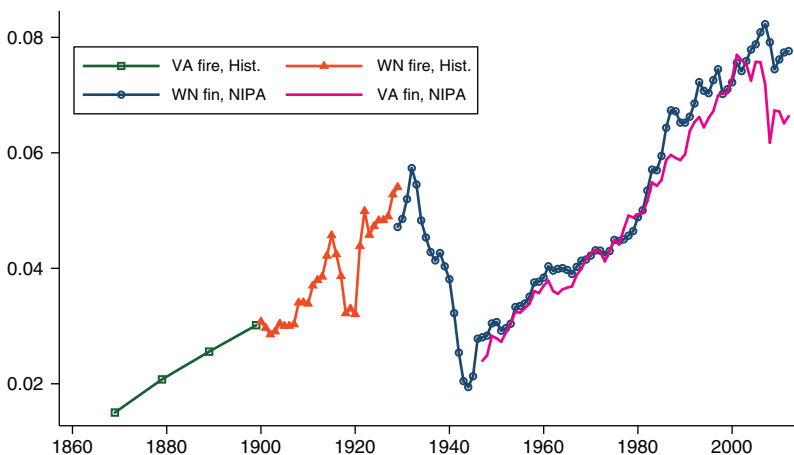


FIGURE 4. INCOME SHARE OF FINANCE INDUSTRY

Notes: VA is value added, WN is compensation of employees, fin means finance and insurance, fire means finance, insurance, and real estate. For NIPA, the data source is the BEA, and for “Hist.” the source is the Historical Statistics of the United States.

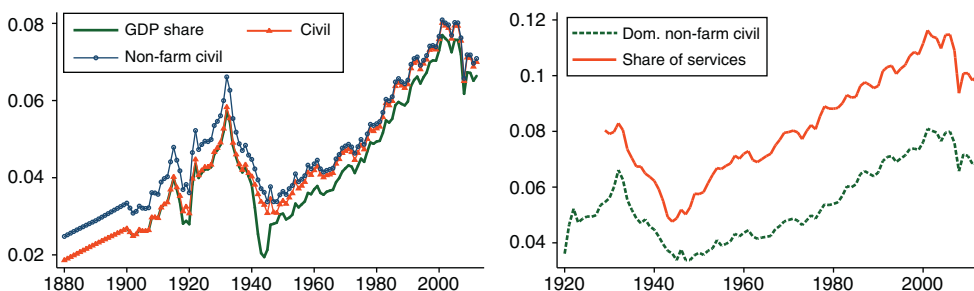


FIGURE 5. INCOME SHARE OF FINANCE (Alternative measures)

Notes: GDP share is the income of the finance industry divided by GDP, constructed from the series in Figure 4. “No Civil Defense” uses GDP minus defense spending, and “No Farm No Civil Defense” uses non-farm GDP minus defense spending. Domestic shares excludes net exports of finance and insurance companies. Share of services uses the BEA definition of services.

defense spending. The second issue is the decline in farming. Since modern finance is related to trade and industrial development, it is also useful to estimate the share of finance in non-farm GDP.

The left panel of Figure 5 presents the finance share of non-defense GDP, and of non-farm, non-defense GDP (or compensation, as explained above). Both adjustments make the series more stationary. In particular, using non-defense GDP removes the spurious temporary drop in the unadjusted series during WWII.

I use the defense-adjusted share as my main measure. The share of finance starts just below 2 percent in 1880. It reaches a first peak of almost 6 percent of GDP in 1932. Note that this peak occurs during the Great Depression, not in 1929. Between 1929 and 1932 nominal GDP shrinks, but the need to deal with rising default rates

and to restructure corporate and household balance sheets keeps financiers busy. Similarly, the post-war peak occurs not in 2007 but in 2010, just below 9 percent of non-defense GDP.

Other Services.—Is finance different from other service industries? Yes. The right panel of Figure 5 also plots the share of finance in service GDP. It is mechanically higher than with total GDP, but the pattern is the same (the other fast growing service industry is health care, but it does not share the U-shaped evolution of finance from 1927 to 2009).

Globalization.—Figure 4 shows finance income divided by US GDP. This might not be appropriate if financial firms export some of their services abroad. It turns out, however, that globalization does not account for the evolution of the finance income share. There are two ways to show this point.

The right panel of Figure 5 displays the ratio of *domestic* finance income to (non-defense) GDP. Domestic income is defined as income minus net exports of financial services. The figure is almost identical to the previous one. The reason is that the United States, unlike the United Kingdom for instance, is not a large exporter of financial services. According to IMF statistics, in 2004, the UK financial services trade balance was +\$37.4 billion while the US balance was −\$2.3 billion: the United States was actually a net importer. In 2005, the UK balance was +\$34.9 billion, and the US balance was +\$1.1 billion. In all case, the adjustments are small.

The timing of globalization also cannot explain the evolution of the US financial sector. Estevadeordal, Frantz, and Taylor (2003) show that the period 1870–1913 marks the birth of the first era of trade globalization (measured by the ratio of trade to output) and the period 1914–1939 its end. The period between 1918 and 1930, however, is the first large scale increase in the size of the finance industry, precisely as globalization recedes. For the more recent period, Obstfeld and Taylor (2002) and Bekaert, Harvey, and Lumsdaine (2002) show that financial globalization happens relatively late in the 1990s, while Figure 1 shows that the growth of the financial sector accelerates around 1980.

II. Quantity of Intermediated Assets

I measure the quantity of intermediated financial assets as follows:

$$(2) \quad q_t \equiv b_{c,t} + m_t + k_t,$$

where $b_{c,t}$ is consumer credit outstanding, m_t are holdings of liquid assets, and k_t is the value of intermediated corporate assets (for the nonfinancial sector). The measurement principle is to take into account the instruments on the balance sheets of end users, households, and nonfinancial firms. This is the correct way to do the accounting, rather than looking at financial intermediaries' balance sheets which reflects (in part) activities within finance itself.

Equation (2) is consistent with a model where it costs the same to extend \$1 of consumer credit, \$1 of business credit (or equity), or to create \$1 of liquidity. That these costs are the same is far from obvious but is in fact consistent with

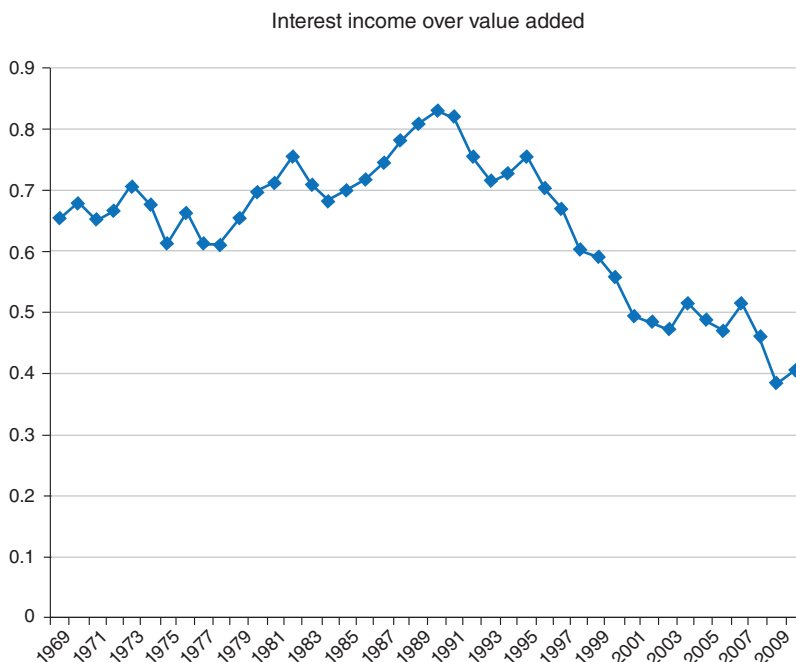


FIGURE 6. COMPARISON WITH INCOME MEASURE IN MEHRA, PIGUILLEM, AND PRESCOTT (2011)

Note: Net interest income as defined in Mehra, Piguillem, and Prescott (2011) divided by value added of financial intermediaries as defined in this paper.

microeconomic evidence available for the more recent part of the sample. These assumptions as well as the underlying model are discussed in the online Appendix. I maintain for now the assumption that the relative costs of various types of intermediation remain constant over time and that the composition of borrowers remains constant. I relax this assumption in Section III. In the remainder of this section, I construct empirical proxies for b_c , m , and k .

A. Debt and Equity

Figure 7 presents credit liabilities of farms, households, and the business sector (corporate and non-corporate). These include all bank loans, consumer credit, mortgages, bonds, etc. The first point to takeaway is the good match between the various sources. As with the income share above, this allows us to extend the series in the past. Two features stand out. First, the nonfinancial business credit market is not as deep even today as it was in the 1920s. Second, household debt has grown significantly over the post-war period.¹⁰

¹⁰I have also constructed credit liabilities of financial firms. Financial firms have recently become major issuers of debt. Banks used to fund themselves with deposits and equity, and almost no long-term debt. Today they issue a lot of long-term debt. Note that it is critical to separate financial and nonfinancial issuers. What should count as output for the finance industry are only issuances by nonfinancial firms.

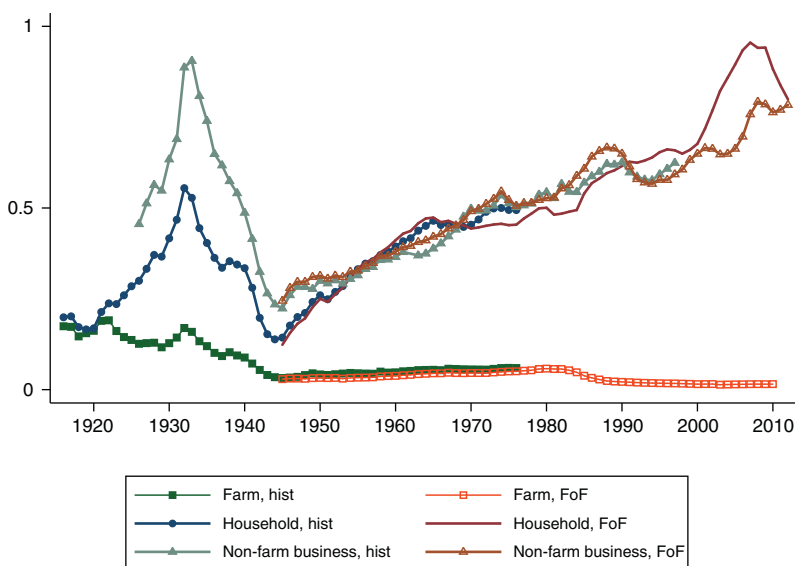


FIGURE 7. DEBT OVER GDP

Notes: FoF is Flow of Funds, Hist is Historical Statistics of the United States. Business includes non-farm corporate and non-corporate debt.

To extend the credit series before 1920, I use data on home mortgages provided by Schularick and Taylor (2012). I also use the balance sheets of financial firms. I measure assets on the balance sheets of commercial banks, mutual banks, savings and loans, federal reserve banks, brokers, and life insurance companies. I define total assets as the sum of assets of all these financial firms over GDP. I use this series to extend the total nonfinancial debt series (households and non-corporates, farms, corporates, government). I regress total credit on total assets and use the predicted value to extend the credit series.

The finance industry not only manages existing assets, but it also originates new assets and replaces old ones as they expire. It is therefore useful to consider stocks and flows separately. Figure 8 shows the issuances of corporate bonds by nonfinancial corporations as well as a measure of household credit flows.¹¹ Note that issuances collapse in the 1930s when the debt to GDP ratio peaks, in part because of deflation. There is thus a difference of timing between measures of output based on flows (issuances) versus levels (outstanding). Figure 8 also shows a measure of household debt issuance.

I use three measures of equity intermediation: total market value over GDP, initial public offering (IPO) proceeds over GDP, and gross (nonfinancial) equity offerings over GDP.¹² Figure 9 shows that gross equity flows were high in the early part of

¹¹ When I do not have a separate measure of flows, I assume a runoff rate consistent with the average ratio of flow to level, and I create the flow measure from the level series. Details are in the online Appendix.

¹² Why use the market value of equity when thinking about intermediation? First, the rise in market value could be driven by improvements in financial intermediation. Two prime examples are risk management with financial derivatives (discussed in the online Appendix) and lower costs of participation in the equity market. Improvement in financial intermediation would lead to higher market value of equity. Clearly, in this case, the measure of unit cost

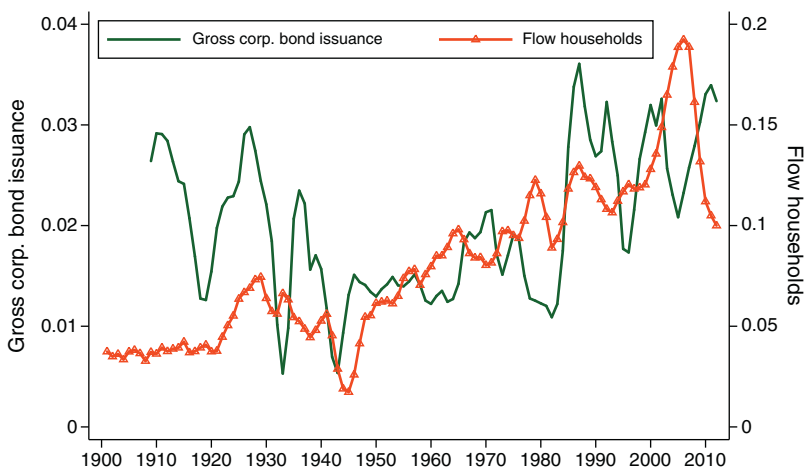


FIGURE 8. DEBT FLOWS OVER GDP

Notes: Gross issuance of corporate bonds is a three-year centered moving average of gross issuances of bonds by nonfinancial firms, from Baker and Wurgler (2000). Household issuance is based on the flow of funds and the historical statistics of the United States.

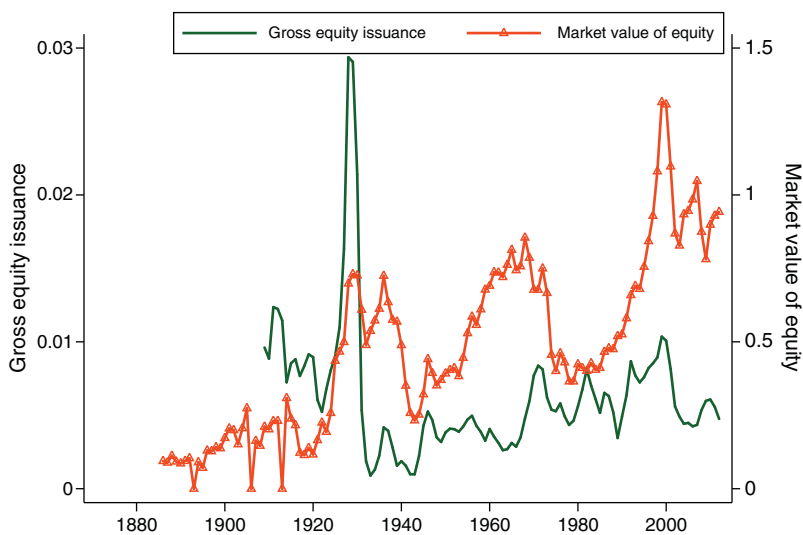


FIGURE 9. EQUITY VALUE AND GROSS ISSUANCE OVER GDP

Notes: Market value of nonfinancial corporate firms from the flow of funds and from the Center for Research on Security Prices (CRSP). Gross equity issuance is a three-year centered moving average of gross issuances of stocks by nonfinancial firms, from Baker and Wurgler (2000).

would be correct only if equity is measured at market value. Another reason for using market values is that book values miss a lot of intangible investments. On the other hand, changes in market values may reflect factors that are not directly related to financial intermediation, such as changes in household risk aversion or bubbles. This then begs the question of what is the “production function” of asset management services. The evidence in Greenwood and Scharfstein (2013) is consistent with a constant fee in the aggregate, even though individual fees might have decreased. Gennaioli, Shleifer, and Vishny (2014a) discuss this issue in details. Finally, notice that in the extended model of Section III, it is only the asset management fee that is proportional. The monitoring cost is not.

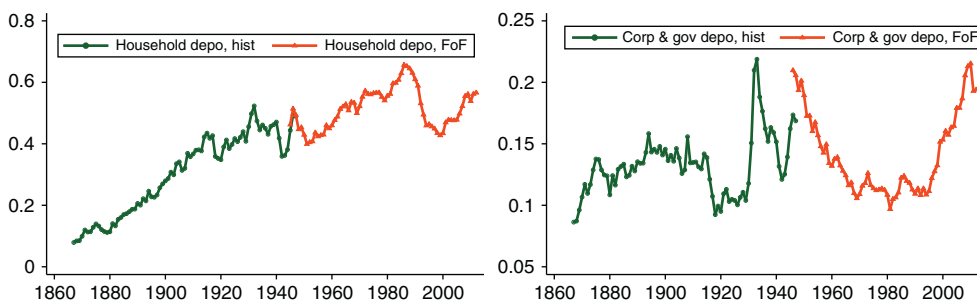


FIGURE 10. LIQUIDITY: DEPOS AND REPOS OVER GDP

Note: Deposits, Repurchase Agreements, and Money Market Mutual Funds shares over GDP.

Sources: Historical Statistics of the United States and Flow of Funds.

the sample. The market value of equity, on the other hand, is higher in the post-war period. The IPO series will allow me to implement quality-adjustments in the next section.

B. Money and Liquidity

In addition to credit (on the asset side of banks), households, firms, and local governments benefit from payment and liquidity services (on the liability side of banks and money market funds). For households, I use total currency and deposits, including money market fund shares, held by households and nonprofit organizations. The left panel of Figure 10 shows the evolution of this variable.

An important element to take into account in the measurement of liquidity provision is the rise of the shadow banking system. Gorton, Lewellen, and Metrick (2012) argue that a significant share recent activities in the financial sector was aimed at creating risk free assets with money-like features. For firms (incorporated or not), I follow Gorton, Lewellen, and Metrick (2012), and I treat repos as shadow deposits. The series is thus the sum of checkable deposits and currency, time and savings deposits, money markets mutual funds shares, and repos (by nonfinancial firms).¹³

C. Aggregation

If we could observe the income flows $y_{i,t}^f$, associated with the three fundamental sources of revenues $i = b, m, k$, we would simply compute the unit cost as, for instance: $\psi_{c,t} = \frac{y_{c,t}^f}{b_{c,t}}$, where $y_{c,t}^f$ would be the income generated by credit intermediation for consumers. Unfortunately, there is no satisfactory way to link a particular income to a particular activity, especially over long periods of time.¹⁴

¹³I have experimented with an adjustment for the fact that deposit insurance provided by the government makes it cheaper for private agents to create deposits. The adjustments seem rather arbitrary and did not make a significant difference so I dropped it. But more quantitative work would clearly be needed here.

¹⁴There is an empirical problem and a conceptual problem. Empirically, our data is organized by industry (e.g., securities, credit intermediation), not by function and even less by end-user. Even obtaining detailed measures of gross output is challenging. See Greenwood and Scharfstein (2013) for an enlightening discussion. But this is not only an issue of accounting. Even if we had all the data imaginable, we would still need to decide how to allocate

This precludes a direct estimation of the $\psi_{i,t}$ s. We only observe the total income of the finance industry, y_t^f described in Section I. This is why I assume that the relative costs remain constant over time. The online Appendix shows how they can be estimated and that, in fact, they are close to one.

M&As.—An important activity of financial intermediaries is advising on mergers and acquisitions (M&As). Rhodes-Kropf and Robinson (2008) show that M&As differ from other types of investment and require specific search efforts. From 1980 to 2010, I use data from Securities Data Company (SDC) and Bloomberg to compute the value of merger deals. I then use historical data from Jovanovic and Rousseau (2005) to extend the series back to 1890. The next step is to apply the proper weight to the M&A series. M&A fees typically range from 1 percent for large deals to 4 percent for smaller ones. I assume that merger fees are 2 percent of the volume. This assumption is probably a bit higher than the weighted average fee, but there are also probably some ancillary activities associated with mergers and for which the finance industry is compensated.

Flows and Stocks of Intermediated Assets.—I construct two measures, one for the flow of new intermediation, one for the stock of outstanding intermediated assets.

Some activities are more naturally linked to flows (screening, IPO fees, etc.), and some are more naturally linked to stocks (debt restructuring, asset management, etc.). The stock measure is simply the sum of outstanding values

$$(3) \quad q_t^{level} = b_{c,t}^{level} + b_{k,t}^{level} + e_{k,t}^{level} + m_t.$$

Note that $e_{k,t}^{level}$ is the *market* value of equity, as discussed earlier. For the flow measure, I also add up the values of new issuances, but I take into account the fact that underwriting fees are higher for equity than for debt (see Altinkiliç and Hansen 2000 and the online Appendix for details),

$$(4) \quad q_t^{flow} = b_{c,t}^{flow} + b_{k,t}^{flow} + 3.5e_{k,t}^{flow} + M\&A_t.$$

It corresponds to gross issuances of debt and equity, plus the value of mergers and acquisitions.¹⁵ Note that the liquidity measure is only a level measure, and that the M&A measure is only a flow measure. Finally, the total measure of intermediated assets is

$$(5) \quad q_t = q_t^{flow} + q_t^{level}.$$

costs among many shared activities: hedging and risk management, trading, over-head labor, etc. And financial tasks are deeply intertwined. Insurance companies and pension funds perform their own independent credit analysis. Banks act as market makers. Investment banks behave as hedge funds. In addition, the mapping from industry to tasks has changed over time with the development of the originate and distribute model in banking. Therefore the problem runs even deeper if we want to make long-run comparisons.

¹⁵In theory, I would also need to take into account the debt of the government. The issue is which weight to apply. Government debt is risk-free and liquid, and it might actually help the functioning of financial markets and justify a negative weight (Krishnamurthy and Vissing-Jorgensen 2012; Greenwood, Hanson, and Stein 2011). But any long-term debt carries duration risk and positive intermediation costs. As a benchmark I set the weight to zero. The results are essentially unchanged if I set the weight to 1/10 instead.

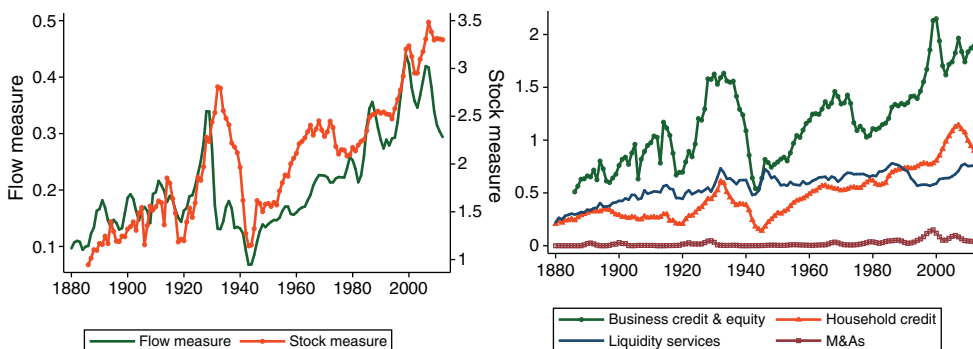


FIGURE 11. INTERMEDIATED ASSETS OVER GDP, STOCKS AND FLOWS

Notes: Left panel shows aggregate measures of stocks and flows of intermediated assets in the United States, as shares of GDP. Right panel shows the sum of stocks and flows sorted by end-user.

The aggregate flow and stock measures (q_t^{flow} , q_t^{level}) are displayed in the left panel of Figure 11. The flow measure is an order of magnitude smaller than the stock measure. The flow measure collapses quickly during the Great Depression while the level measure peaks later and is exacerbated by deflation. A similar pattern emerges during the Great Recession. Overall, the stock measure increases more in recent years, driven by the market value of corporate equity and by the size of the household debt market.

The right panel of Figure 11 presents the total measures corresponding to four broad functions discussed earlier: credit and equity intermediation services to firms, credit intermediation services to households, liquidity services, and M&A activities. It is clear from Figure 11 that the intermediation series for firms and households are the most volatile ones. There is also a significant increase in liquidity services in the 2000s. M&As play some role mostly in the 1990s. By construction, the sum of the two series in the left panel of Figure 11 is the same as the sum of the four series in the right panel, and is equal to the measure of intermediated assets, q , in Figure 2.

D. Evidence of Constant Returns to Scale

Figure 3 shows the raw estimate of the cost of financial intermediation ψ_t , defined as income divided by intermediated assets. For income, I use domestic income, i.e., income minus net exports, as explained in Section I. Before discussing quality adjustments in the next section, I present evidence of constant returns to scale in financial intermediation.

An important assumption of the model is that financial services are produced under constant returns to scale. Figure 12 presents evidence consistent with this assumption. It uses the period 1947–1973, for two reasons. First, the post-war data is the most reliable, and stopping in 1973 allows me to exclude major oil shocks, inflation, and other factors that might create short-term noise in my estimates. Second, as I will discuss shortly, quality adjustments are less important over this period than either before or after. Since these adjustments are difficult to implement, it is more convincing to first present the evidence without them.

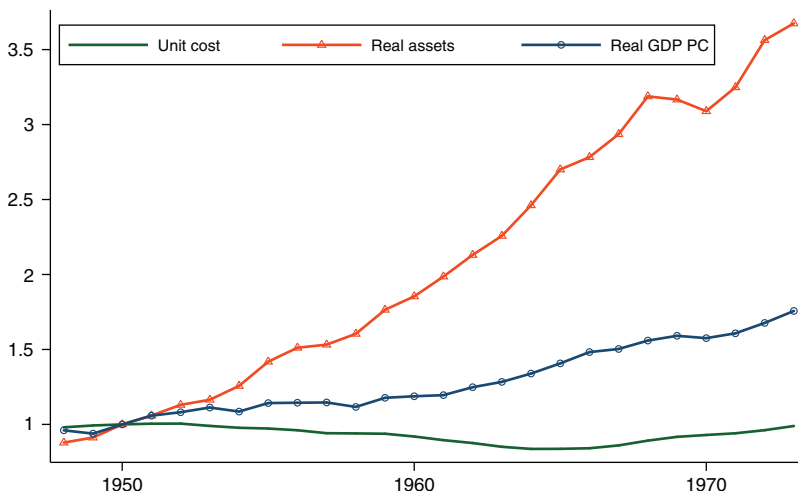


FIGURE 12. CONSTANT RETURNS TO SCALE

Notes: Unit cost of financial intermediation, real intermediated assets, and real GDP per capita. Series normalized to one in 1950.

From 1947 to 1973, real GDP per capita increases by 80 and real financial assets by 250 percent (measured in constant dollars), but my estimate of the unit cost of intermediation remains fairly constant (all series are presented as ratios to their values in 1950). By 1970 people are a lot richer, financial markets are a lot larger, but the unit cost is exactly the same as in 1950. This provides clear evidence that the production of financial services has constant returns to scale.

III. Quality Adjustments

The quantities of intermediation should be adjusted for the difficulty of monitoring/screening borrowers, otherwise the unit cost measure could register spurious changes in intermediation efficiency. These adjustments require a model. The model economy consists of households, a nonfinancial business sector, and a financial intermediation sector. In the model, the finance industry provides three types of services to households and firms: liquidity, monitoring, and asset management. Households hold the corporate capital stock via intermediaries. In addition, households borrow and lend from each other. The key point of the model is that households and firms are heterogeneous in their intermediation intensities. Some borrowers/issuers require more screening and monitoring than others.

A. Corporate Finance

The homogeneous borrower model used earlier is a useful benchmark, but it fails to capture some important features of corporate finance. To give just one example, corporate finance involves issuing commercial paper for blue chip companies as well as raising equity for high-technology start-ups. The monitoring requirements

per dollar intermediated are clearly different in these two activities. Measurement problems arise when the mix of high- and low-quality borrowers changes over time. Constant heterogeneity does not pose a problem: it amounts to a simple rescaling of the unit cost in Figure 3. Changes in the share of hard-to-monitor projects, however, present a challenge.

Let us therefore consider a simple moral hazard model with heterogeneous firms. If a firm hires n workers it produces $f(n)$ units of output, where f is increasing and concave. Firms choose employment to maximize (detrended) net income $\pi(w) \equiv \max_n f(n) - wn$. There are two types of firms, l and h , that differ in their cash on hand x (equivalently in their retained earnings or their pledgeable collateral). I assume that $x^l < x^h$ and I refer to l -firms as low cash firms. There is an exogenous potential supply k_h of h -firms and free entry of l -firms.¹⁶ To capture financial intermediation in a tractable way, I assume that capital can be diverted. The online Appendix describes the details of moral hazard and endogenous monitoring. The key point is that the model delivers the following monitoring demand function

$$\mu(x) = r + \delta + \varphi - \pi(w) + (1 + r)(\xi - x),$$

where ξ is the fraction of capital that can be diverted if there is no monitoring and φ is a proportional intermediation cost, akin to asset management fees. The function $\mu(x)$ measures the quantity of intermediation services required for a firm with cash on hand x . Firms with high values of x require less monitoring than firms with low values of x . The unit cost of monitoring is ζ_t and the income received by intermediaries for their monitoring activity is $\zeta_t \bar{\mu}_t$, where aggregate monitoring is

$$\bar{\mu}_t \equiv \mu_h + (1 + r)(x_h - x_l)s_t,$$

and $s_t \equiv \frac{k_{l,t}}{k_{l,t} + k_{h,t}}$ is the share of low cash firms in aggregate investment. The total income for corporate intermediation services is

$$y_{k,t}^f = \varphi_t k_t + \zeta_t \bar{\mu}_t.$$

Similarly, external finance (the quantity of monitored assets) is $\bar{b}_k = 1 - x^h + (x_h - x_l)s$. Note that the unit cost of external finance $\zeta_t \frac{\bar{\mu}_t}{\bar{b}_t}$ depends on the intensity of monitoring $\bar{\mu}/\bar{b}$, which changes with the share of low cash firms s . The parameter of interest is ζ_t which captures the true efficiency of financial intermediation. To recover ζ_t , I need to estimate $\bar{\mu}/\bar{b}$.

¹⁶Let k_t be the (endogenous) number of active firms, and let n_t be employment per firm (so aggregate employment is $\bar{n}_t = k_t n_t$). The number k_h captures the extent to which investment opportunities occur in established companies. I assume that it is given by technology, and indeed, the data supports the view that large changes in k_h are driven by large scale technological change (electricity, information technology). Note that the number of low cash firms $k_l - k_h$ is endogenous, and in particular, highly dependent on financial intermediation. So the way the model is going to interpret the 1990s is that established firms were not the ones able to promote the IT revolution. Instead it had to be younger firms, that are cash poor and therefore more dependent on financial intermediation. To the extent that we actually observe a large entry of young firms, the model will infer that financial intermediation must have been relatively efficient.

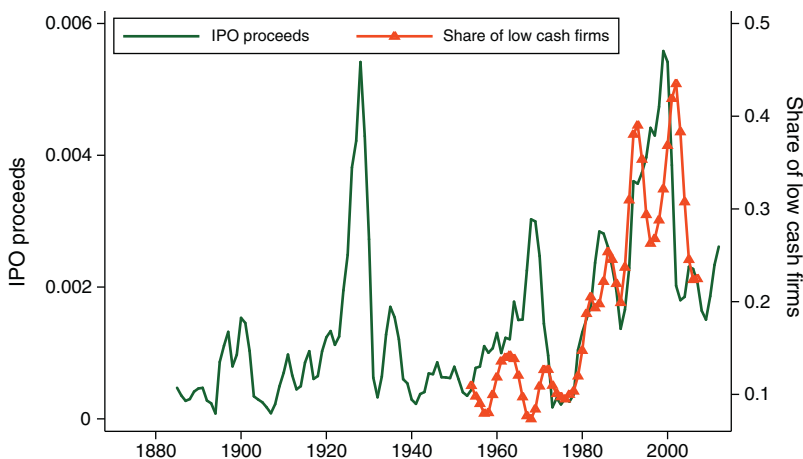


FIGURE 13. INPUTS FOR CORPORATE FINANCE QUALITY ADJUSTMENTS

Notes: The share of low cash firms is from Philippon (2008). The IPO series is a three-year centered moving average of IPO proceeds over GDP.

Source: Jovanovic and Rousseau (2001) and Ritter (2011).

Philippon (2008) uses Compustat to construct an empirical proxy for s_t , namely the share of aggregate investment that is done by firms that must borrow more than three-quarters of their capital spending. The measure is displayed Figure 13. Following Eisfeldt and Rampini (2006), I have also computed measures of investment that include capital reallocation by adding acquisitions minus sales of used capital for each firm. All these measures are similar and suggest that the intensity of corporate finance was higher in the 1980s and 1990s than in the 1960s. Since these measures are based on Compustat data, they are available only from 1950 onward (at best). Figure 13 also shows IPO proceeds, based on the work of Jovanovic and Rousseau (2001) and Ritter (2011). The two series are highly correlated in the post-war period, and I use the IPO series to extrapolate the low cash share series before 1950, using a simple linear regression of one variable on the other. As argued by Jovanovic and Rousseau (2001), the IPO market of the 1920s was remarkably active, even compared to the one of the 1990s: IPO firms were of similar ages, and the proceeds (as share of GDP) were comparable.

B. Household Finance

On a per-dollar basis, it is more expensive to lend to poor households than to wealthy ones, and relatively poor households have gained access to credit in recent years.¹⁷ To capture this idea, I assume that there is a continuum households and that there is a fixed cost to borrowing κ , in addition to the marginal cost φ . Income

¹⁷Using the Survey of Consumer Finances, Moore and Palumbo (2010) document that between 1989 and 2007 the fraction of households with positive debt balances increases from 72 to 77 percent. This increase is concentrated at the bottom of the income distribution. For households in the 0–40 percentiles of income, the fraction with some debt outstanding goes from 53 to 61 percent between 1989 and 2007. In the mortgage market, Mayer and Pence

inequality among households is captured by the labor endowment η . The model is described in details in the online Appendix. The model features both an extensive margin (participation of households in the credit market) and an intensive margin (how much each household borrows). The extensive margin is characterized by the cutoff $\hat{\eta}$ such that only households with income above $\hat{\eta}$ use the credit market. The fraction of households (of a given generation) who have access to credit is therefore $1 - F(\hat{\eta})$ where F is the c.d.f. of η .

The aggregate stock of household debt, relative to labor income w , is

$$\frac{\bar{b}_c}{w} = \frac{1 + \gamma}{2 + r + \gamma} \int_{\eta > \hat{\eta}} \left((\lambda - (1 - \varphi)^{-1}) \eta - \kappa \right) dF(\eta),$$

where γ is the rate of growth of the economy, and λ is the slope of life-cycle earnings, which determines the desire to borrow in order to smooth consumption. The income the finance industry receives from consumers credit is

$$y_c^f = \varphi \bar{b}_c + \kappa w (1 - F(\hat{\eta})).$$

C. Calibrated Model

The last step is to calibrate the model and construct the required quality adjustments. I rely as much as possible on micro-evidence to pin down the parameters of the model. I can then reduce the number of unknown parameters to seven, which I estimate using eight moments, so the model is slightly over-identified. An important variable is the income of the finance industry,

$$(6) \quad y^f = \varphi(k + \bar{b}_c) + \zeta \bar{\mu}(s) + \kappa w (1 - F(\hat{\eta})) + \psi_m m.$$

I have assumed that the linear cost (asset management) φ is the same for corporate and household finance. The parameters s and $\hat{\eta}$ capture changes in the characteristics of borrowers. I use 1989 as a reference year because of data availability. The details of the calibration are presented in the online Appendix. The model matches the size of the various markets, the fraction of low cash firms, the participation rate of households in credit markets, and the income of the finance industry, all measured in 1989. The implied parameters are reasonable. For instance, I estimate a fixed borrowing cost κ of 2 percent. In the model, the finance industry earns 1.35 percent of GDP from liquidity, 2.08 from household credit, and 2.37 from business intermediation, for a total of 5.8 percent of GDP.

The calibrated version can then be used to understand the qualitative properties of the model and the biases that could arise in the measurement of financial intermediation. There are two types of biases. The first type of bias is that, holding intermediation technology constant, changes in the characteristics of borrowers can affect the measured unit cost of intermediation. The second type of bias comes from changes in the intermediation technology itself.

(2008) show that subprime originations account for 15–20 percent of all Home Mortgage Disclosure Act (HMDA) originations in 2005.

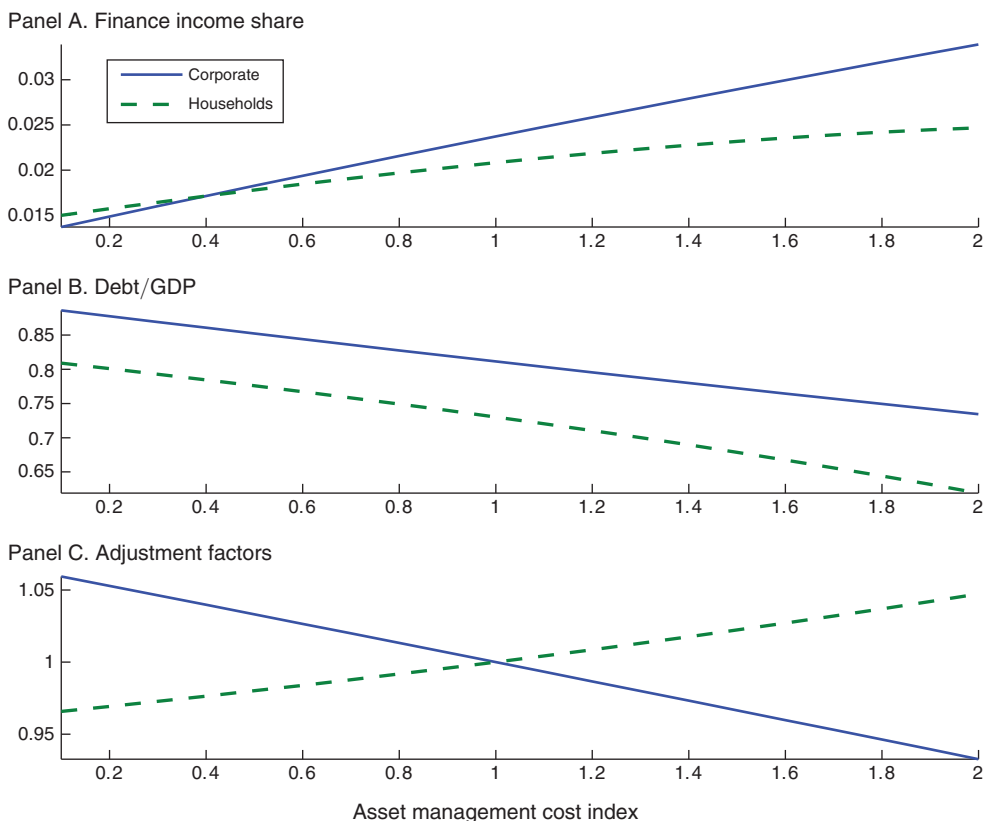


FIGURE 14. QUALITY ADJUSTMENTS FOR ASSET MANAGEMENT COSTS

Notes: Comparative statics using the calibrated model (Table 4 of online Appendix). The horizontal axis is φ/φ_0 where φ_0 is calibrated to 1 percent in 1989. Panel A presents separately the income from corporate finance services and the income from household finance services. Panel B shows firm’s external finance and household debt. Panel C shows the adjustment to the particular class of intermediated assets needed to remove the bias in the measurement of intermediation costs. When the adjustment is above 1, the unit cost is overestimated and the quantity of assets must be scaled up to obtain the correct estimate.

Figure 14 studies the impact of changes in the unit cost of asset management φ , which is calibrated to 1 percent in 1989. An increase in the cost of asset management increases the finance share of GDP (panel A), and decreases the size of the credit market (panel B), as expected. Note that the model, unlike the data, allows me to separate the income received from corporate finance and from household finance. The question is whether the unit costs $\psi_k = y_k^f/k$ and $\psi_c = y_c^f/\bar{b}_c$ correctly capture the changes in φ .¹⁸ Panel C shows that the answer is “almost.” Even for very large

¹⁸The model gives a mapping Q from the parameters $\chi = (\varphi, \kappa, k_h, \dots)$ to the equilibrium objects $(k, s, \bar{b}_k, \bar{b}_c, \hat{\eta}, \dots) = Q(\chi)$. The income of the finance industry y^f in equation (6) depends on $Q(\chi)$ and on the intermediation technology: I write it as $y^f = Y(\chi, Q)$. In the model, this income can be further decomposed into the components coming from different types of intermediation. We have the empirical measure q from equations (4), (3), and (5) which is simply a linear combination of the elements in Q . The measured unit cost is $\psi = y^f/q$. Starting from a benchmark equilibrium (χ_0, Q_0) , suppose that we change one parameter, says φ , so that $\chi = (\zeta_0, k_h, \varphi, \kappa_0, \dots)$. Using the model, we can compute Q, q , and y^f , and the measured unit cost $\psi = y^f/q$. The measured change in unit cost is ψ/ψ_0 . If the model is linear in the parameter of interest as in Section II, then we have $\psi/\psi_0 = \varphi/\varphi_0$. But in general the model is not linear. To fix intuition, suppose that $\psi/\psi_0 > 1$ so we looking at a perceived increase in

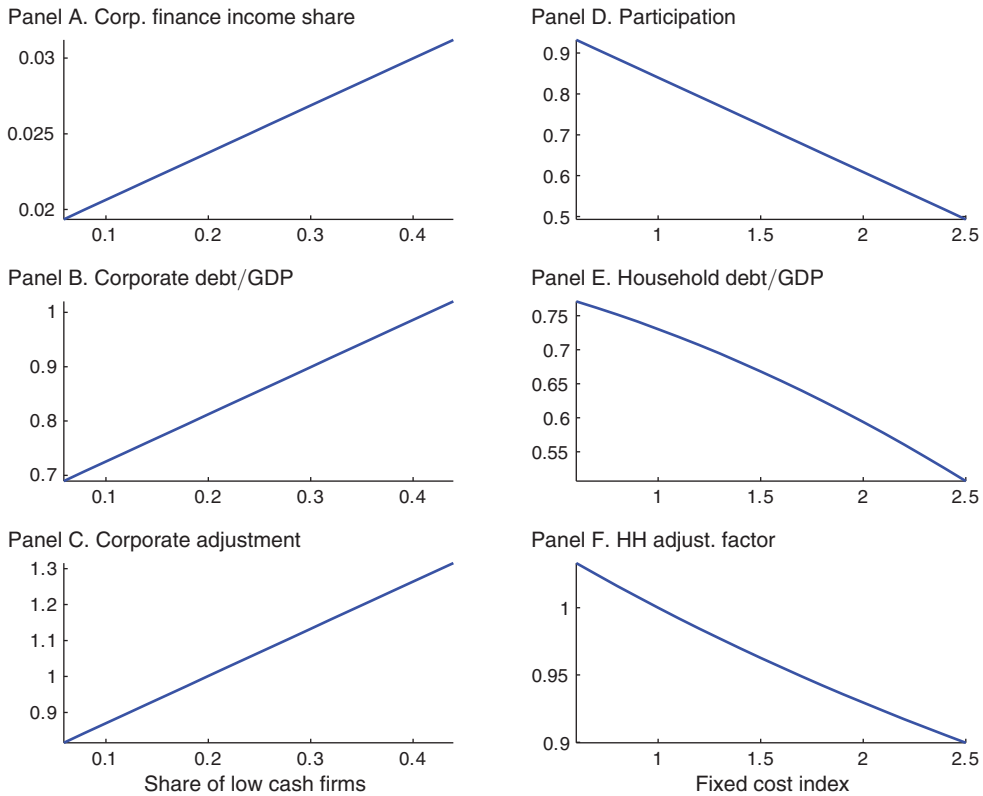


FIGURE 15. QUALITY ADJUSTMENTS FOR FIRMS CHARACTERISTICS AND PARTICIPATION COSTS

Notes: Comparative statics using the calibrated model (Table 4 of online Appendix). Adjustment is the scaling up or down of the particular class of intermediated assets needed to remove the bias in the measurement of the unit cost. The forcing variables are k_h in panels A, B, C, and κ in panels D, E, F.

changes in φ (from 0 to twice the benchmark value), the bias barely exceeds 5 percent. In a sense, this is not surprising because we are only changing the linear part of the model in this experiment. When φ goes down, more potential borrowers actually borrow, and each borrower borrows more. This does not create a significant bias.

Figure 15, on the other hand, studies the highly nonlinear part of the model, by changing the share of low cash firms and the fixed cost of participation for households. Panels A, B, and C focus on changes in the composition of firms. The exogenous forcing variable is the number of cash-rich firms k_h . Note that the figure shows the response of the various variables as a function of the observed share of low cash firm s , which is itself an endogenous variable. The reason is that s is observable in the data and will be used to make the quality adjustment. The benchmark model is calibrated using a share of low cash firms of 20 percent (in 1989). When this

the unit cost. To adjust this measure, I define \hat{y}_0^f as counterfactual income if we wanted to obtain Q with the initial technology ζ_0 , in other words $\hat{y}_0^f = Y(\chi_0, Q)$. The conceptually correct change in the unit cost is y^f/\hat{y}_0^f since by construction Q is unchanged. If $\psi/\psi_0 > y^f/\hat{y}_0^f$, then the empirical measure overestimates the change in unit cost. Then the adjustment is defined as $\psi\hat{y}_0^f/\psi_0y^f$. This adjustment has the property that if I use it to artificially scale up q , I recover the correct value for $\psi/\psi_0 = y^f/\hat{y}_0^f$. This adjustment can be applied to the entire amount of intermediated assets, or to particular classes, such as b_c, b_k , etc.

share increases, monitoring costs and external finance both increase (panels A, B). Monitoring intensity increases, and this creates a measurement bias in the sense that the perceived unit cost increases. If the share is 40 percent, the model says that external corporate finance should be scaled up by roughly 25 percent in order to remove the induced bias in the measurement of the unit cost.

Panels D, E, and F in Figure 15 focus on changes in the availability of household credit induced by exogenous changes in the fixed cost κ . When κ increases, some relatively poor households are priced out and participation in the credit market falls (D). The model is calibrated to a participation of 84 percent and a household debt to GDP ratio of 73 percent in 1989. If κ doubles, the participation rate and the debt/GDP ratio drop to approximately 60 percent. The participation rate drops more than the debt/GDP ratio because rich households still borrow, and they typically borrow more (when young) than poor households. These nonlinear composition effects create again a significant bias.

D. Adjusted Unit Cost

The goal of this section is to use the calibrated model presented above to adjust the asset series of Figures 11. The first step is to choose which adjustments to make. I take away from Figure 14 that changes in the proportional cost φ are unlikely to create significant biases. I therefore focus on the other parameters. At the firm level, the choice is fairly obvious: Figure 13 shows that the share of low cash firms changes a lot over time, and panel C of Figure 15 shows that this can create large biases. The implementation is straightforward: plug in the observed value of s , read the adjustment factor in panel C of Figure 15, and multiply the empirical series for \bar{b}_k by this factor. The implied series is “Firm adj.” in Figure 16. As expected, the adjustment is quantitatively important in the 1920s and in the 1980s and 1990s, which correspond to waves of innovation driven by new technologies.

Biases in the household debt market are likely to come from changes in household participation. The adjustment is difficult because I do not have a long time series for the participation rate of households. Since the goal of this section is to assess measurement biases, I will look for the maximal adjustment by assuming that changes in the household debt to GDP ratios are driven by changes in the fixed cost κ .¹⁹

The series “Firm and HH adj.” in Figure 16 shows the output measures with quality adjustments for both corporate and household finance. Adjustments to consumer credit matters mostly after 1970. As argued earlier the quality adjustments are small between 1947 and 1973, which makes it an ideal period to test the constant returns to scale assumption. The adjustment is large in the recent part of the sample. After

¹⁹There is prima facie evidence of technological change in the intermediation technology (e.g., credit scoring) that has made it easier for poor households to obtain credit. So we know that this account for *some* of the evolution of the household debt market, but we do not know precisely how much. I am going to interpret the historical time series as if the growth in consumer credit mostly reflects improvements in intermediation. I only impose the constraint that the predicted participation rate cannot exceed 100 percent. This constraint binds in the model in the years 2000s, which is consistent with the view that household debt growth was linked to house prices for households who already had access to credit. I have also considered the implications of changes in inequality, but I have found that these are unlikely to create significant biases. Changes in inequality typically change the debt/GDP ratio and the finance income share, but the quantitative experiments suggest that the unit cost is not severely biased.

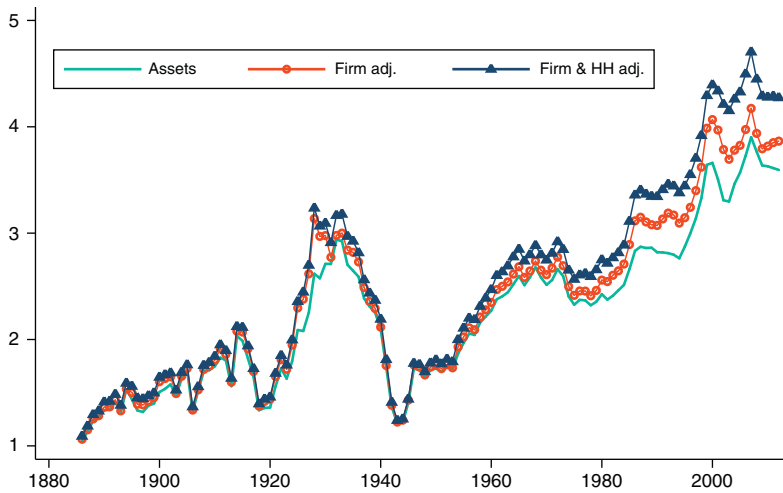


FIGURE 16. QUALITY ADJUSTED INTERMEDIATED ASSETS

Notes: Adjustments are computed using the calibrated model. Firm adj. takes into account changes in the fraction of low cash firms, using the series in Figure 13 and the adjustment function in panel C of Figure 15. Firm and HH adj. assumes in addition that the long-term growth in household finance is driven by expanding access to credit (lower κ).

1990 the unadjusted measure of business intermediation underestimates intermediation by about 25 percent.

Table 1 and Figures 16 and 17 are the main contributions of the paper. They bring together the historical/empirical work of Sections I and II, and the theoretical/quantitative work of Section III (an adjustment for non-life insurance services is discussed below). Figure 17 shows the unit cost of financial intermediation, defined as income divided by adjusted intermediated assets. There are two main points to take away. The first is that the unit cost ratio is remarkably stable. Recall that we start from series—for income, debt, equity, etc.—that fluctuate a lot over time. But their ratio is stable. The simple unit cost series has a mean of 1.87 percent and a volatility of 23 basis points. Quality adjustments increase the volatility of the assets series but reduce the volatility of the unit cost measure, by about 25 percent. The adjusted series has a standard deviation of only 16 basis points. The second main point is that the unit cost of financial intermediation is about the same today as it was around 1960 and 1900.

E. Insurance Services

The model is designed to account for consumption smoothing that takes place via credit markets. In the model, an improvement in household finance leads to more borrowing and better consumption smoothing.²⁰ Insurance companies, however,

²⁰For instance, Gerardi, Rosen, and Willen (2010) find that the purchase price of a household's home predicts its future income. The link is stronger after 1985, which coincides with important innovations in the mortgage market. The increase in the relationship is more pronounced for households more likely to be credit constrained. The model

TABLE 1—ESTIMATED FINANCIAL INTERMEDIATION

Variable	Observations	Mean	SD	Min	Max
Finance income/GDP	127	0.0422	0.0169	0.0194	0.077
Finance income/GDP w/o ins.	127	0.0359	0.0132	0.0163	0.063
Intermediated assets/GDP	127	2.241	0.726	1.046	3.90
Inter. assets, firm adj.	127	2.369	0.816	1.062	4.17
Inter. assets, firm and HH adj.	127	2.511	0.943	1.091	4.70
Unit cost	127	0.0187	0.00221	0.0145	0.0235
Unit cost, firm adj.	127	0.0178	0.00176	0.0140	0.0222
Unit cost, firm and HH adj.	127	0.0169	0.00165	0.0132	0.0217
Unit cost, adj. w/o ins.	127	0.0146	0.00208	0.0104	0.0204

Note: Data range 1886–2012.

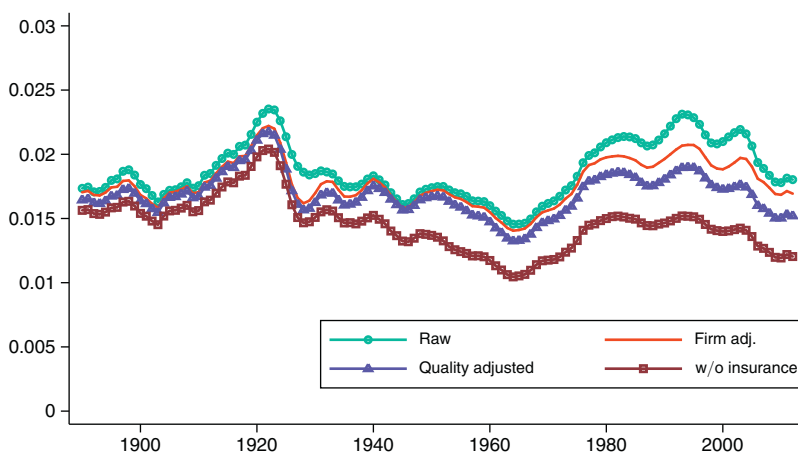


FIGURE 17. QUALITY ADJUSTED UNIT COST OF INTERMEDIATION

Note: Total intermediation costs divided by quality-adjusted intermediated assets.

provide services that are not directly related to intermediation. This is potentially an issue since the income of insurance companies is counted as a cost of intermediation, while the services provided might not be well captured by standard measures of intermediated assets. I therefore attempt a (rough) adjustment by subtracting consumption expenditures on non-life insurance services (health insurance, household insurance, motor vehicle, and other transportation insurance) from the total income of intermediaries. The quantitative significance of this adjustment comes from motor vehicle insurance which grows rapidly in the 1950s and is around one-half

captures correctly measures these effects, and consumption smoothing that entails the creation of credit flows does not create a bias in my estimation. Informal risk sharing, for instance within families, would be enter neither the income side, not the asset side of my calculations, so it should not create a bias either. The overall evidence on risk sharing is mixed. Income inequality has increased dramatically in the United States over the past 30 years. If financial markets improve risk sharing, however, one would expect consumption inequality to increase less than income inequality. This is a controversial issue, but Aguiar and Bils (2011) find that consumption inequality closely tracks income inequality over the period 1980–2007. Therefore it seems difficult to argue that risk sharing among households has improved significantly over time.

of 1 percent of GDP today, and health insurance which grows linearly to reach about 1 percent of GDP. Whether or not these services ought to be included in financial intermediation is debatable. On the one hand, these services differ significantly from banking and traditional intermediation services. On the other hand they *are* financial services linked to the consumption of particular goods, and they certainly affect precautionary savings decisions and therefore the size of the credit market. Removing all of these expenditures is probably an over-adjustment, so the unit cost without (non-life) insurance in Figure 17 should be seen as a lower bound on the true unit cost. The new series suggests a slight downward trend in unit cost until 1970. It does not change the main point regarding the post-war sample: the unit cost is still low in the 1960s, and the discrepancy with the 2000s is at least as large as before.

F. Discussion of the Results

Even after taking into account the various adjustments described above, the unit cost of financial intermediation appears to be as high today as it was around 1900.²¹ This is puzzling. Advances in information technology (IT) should lower the physical transaction costs of buying, pooling, and holding financial assets. Trading costs have indeed decreased (Hasbrouck 2009), but trading volumes have increased even more, and active fund management is expensive.²² French (2008) estimates that investors spend 0.67 percent of asset value trying (in vain on average, by definition) to beat the market. Similarly, Greenwood and Scharfstein (2013) show that, while mutual funds fees have dropped, high fee alternative asset managers have gained market share. The end result is that asset management unit costs have remained roughly constant. The comparison with retail and wholesale trade is instructive. In these sectors Philippon (2012) shows that larger IT investment coincides with lower prices and lower (nominal) GDP shares. In finance, however, exactly the opposite happens: IT investment and the income share are positively related.

A potential explanation is oligopolistic competition but the link between market power and the unit cost of intermediation is not easy to establish. Adding a constant markup of price over marginal cost would not change anything to the trends presented above. The issue is whether market power changes over time. The historical evidence does not seem to support the naïve market power explanation, however. Regulatory barriers to entry have been reduced in banking since the 1970s and yet this is when the unit cost goes up. Conversely, if there is one period where we have a strong presumption that banks had significant market power, this must be the turn of the nineteenth century. But the late Gilded Age is not a period where the unit cost of intermediation is high. The link between market power and the unit cost is therefore rather tenuous and more research is needed on this important topic.

²¹ One should keep in mind that the adjustments are likely to provide lower bounds on the unit cost. Another important point is that I measure equity at market value. In equilibrium, if the cost of holding a diversified portfolio goes down, then the value of the portfolio should go up. My measure attributes the entire secular increase in the price-earnings ratio to an improvement in financial intermediation.

²² Why do people trade so much? Financial economics does not appear to have a good explanation yet. One explanation is overconfidence, as in Odean (1998). Recent work by Glode, Green, and Lowery (2012) and Bolton, Santos, and Scheinkman (2011) explains why some type of informed trading might be excessive. Pagnotta and Philippon (2011) present a model where trading speed can be excessive. In these models, advances in IT do not necessarily improve the efficiency of financial markets.

Another plausible explanation is that my measures might fail to capture the social value of information production in financial markets. This effect is elusive because it can show up as an improvement in TFP with little impact on the aggregate quantity of assets. The only way to test the information production hypothesis is then to estimate directly the informational content of asset prices, as Bai, Philippon, and Savov (2011) attempt to do. This is another area where more research is needed.

IV. Conclusion

I have provided benchmark measures for the aggregate income of the US finance industry, the quantity of intermediated assets, and the unit cost of financial intermediation. The income of the finance industry as a share of GDP fluctuates a lot over time. These fluctuations are mostly driven by equally large fluctuations in the quantity and quality of intermediated assets. The unit cost of financial intermediation represents an annual spread of 1.87 percent on average. The unit cost of intermediation does not seem to have decreased significantly in recent years, despite advances in information technology and despite changes in the organization of the finance industry. The methodology developed in this paper can be used to quantify these evolutions, as well as to compare the cost of financial intermediation across different countries.

REFERENCES

- Acharya, Viral V., Lasse H. Pedersen, Thomas Philippon, and Matthew Richardson. 2009. "Measuring Systemic Risk." <http://pages.stern.nyu.edu/~tphilipp/papers/Systemic.pdf>.
- Admati, Anat R., Peter M. DeMarzo, Martin F. Hellwig, and Paul Pfleiderer. 2011. "Fallacies, Irrelevant Facts, and Myths in the Discussion of Capital Regulation: Why Bank Equity is Not Expensive." <http://www.gsb.stanford.edu/gsb-cmis/gsb-cmis-download-doc/311031>.
- Adrian, Tobias, and Hyun Song Shin. 2014. "Procyclical Leverage and Value-at-Risk." *Review of Financial Studies* 27 (2): 373–403.
- Aguiar, Mark A., and Mark Bilal. 2011. "Has Consumption Inequality Mirrored Income Inequality?" National Bureau of Economic Research Working Paper 16807.
- Altinkılıç, Oya, and Robert S. Hansen. 2000. "Are There Economies of Scale in Underwriting Fees? Evidence of Rising External Financing Costs." *Review of Financial Studies* 13 (1): 191–218.
- Bai, Jennie, Thomas Philippon, and Alexi Savov. 2011. "Have Financial Markets Become More Informative?" <http://pages.stern.nyu.edu/~tphilipp/papers/BaiPhilipponSavov.pdf>.
- Baker, Malcolm, and Jeffrey Wurgler. 2000. "The Equity Share in New Issues and Aggregate Stock Returns." *Journal of Finance* 55 (5): 2219–57.
- Basu, Susanto, Robert Inklaar, and J. Christina Wang. 2011. "The Value of Risk: Measuring the Service Output of US Commercial Banks." *Economic Inquiry* 49 (1): 226–45.
- Bazot, Guillaume. 2013. "Financial Consumption and the Cost of Finance: Measuring Financial Efficiency in Europe (1950–2007)." Paris School of Economics Working Paper 2014–17.
- Beck, Thorsten, Asli Demirgüç-Kunt, and Ross Levine. 2001. "The Financial Structure Database." In *Financial Structure and Economic Growth: A Cross-Country Comparison of Banks, Markets, and Development*, edited by Asli Demirgüç-Kunt and Ross Levine, 17–80. Cambridge, MA: MIT Press.
- Bekaert, Geert, Campbell R. Harvey, and Robin L. Lumsdaine. 2002. "Dating the Integration of World Equity Markets." *Journal of Financial Economics* 65 (2): 203–47.
- Bernanke, Ben S., Mark Gertler, and Simon Gilchrist. 1999. "The Financial Accelerator in a Quantitative Business Cycle Framework." In *Handbook of Macroeconomics*. Vol. 1C, edited by John B. Taylor and Michael Woodford, 1341–93. Amsterdam: Elsevier North Holland.
- Bolton, Patrick, Tano Santos, and Jose A. Scheinkman. 2011. "Cream Skimming in Financial Markets." National Bureau of Economic Research Working Paper 16804.
- Buera, Francisco J., Joseph P. Kaboski, and Yongseok Shin. 2011. "Finance and Development: A Tale of Two Sectors." *American Economic Review* 101 (5): 1964–2002.

- Carter, Susan B., Scott Sigmund Gartner, Michael R. Haines, Alan L. Olmstead, Richard Sutch, Gavin Wright, eds. 2006. *Historical Statistics of the United States, Millennial Edition Online*. Cambridge, MA: Cambridge University Press.
- Christiano, Lawrence, and Daisuke Ikeda. 2011. "Government Policy, Credit Markets and Economic Activity." National Bureau of Economic Research Working Paper 17142.
- Corsetti, Giancarlo, Keith Kuester, André Meier, and Gernot J. Müller. 2013. "Sovereign Risk, Fiscal Policy, and Macroeconomic Stability." *Economic Journal* 123 (566): F99–132.
- Cúrdia, Vasco, and Michael Woodford. 2010. "Conventional and Unconventional Monetary Policy." *Federal Reserve Bank of St Louis Review* 92 (4): 229–64.
- Diamond, Douglas W. 1984. "Financial Intermediation and Delegated Monitoring." *Review of Economic Studies* 51 (3): 393–414.
- Diamond, Douglas W., and Philip H. Dybvig. 1983. "Bank Runs, Deposit Insurance, and Liquidity." *Journal of Political Economy* 91 (3): 401–19.
- Diamond, Douglas W., and Raghuram G. Rajan. 2001. "Liquidity Risk, Liquidity Creation, and Financial Fragility: A Theory of Banking." *Journal of Political Economy* 109 (2): 287–327.
- Eisfeldt, Andrea L., and Adriano A. Rampini. 2006. "Capital Reallocation and Liquidity." *Journal of Monetary Economics* 53 (3): 369–99.
- Estevadeordal, Antoni, Brian Frantz, and Alan M. Taylor. 2003. "The Rise and Fall of World Trade, 1870–1939." *Quarterly Journal of Economics* 118 (2): 359–407.
- French, Kenneth R. 2008. "Presidential Address: The Cost of Active Investing." *Journal of Finance* 63 (4): 1537–73.
- Gennaioli, Nicola, Andrei Shleifer, and Robert W. Vishny. 2013. "A Model of Shadow Banking." *Journal of Finance* 68 (4): 1331–63.
- Gennaioli, Nicola, Andrei Shleifer, and Robert Vishny. 2014a. "Finance and the Preservation of Wealth." *Quarterly Journal of Economics* 129 (3): 1221–54.
- Gennaioli, Nicola, Andrei Shleifer, and Robert Vishny. 2014b. "Money Doctors." *Journal of Finance*. doi: 10.1111/jofi.12188.
- Gerardi, Kristopher S., Harvey S. Rosen, and Paul S. Willen. 2010. "The Impact of Deregulation and Financial Innovation on Consumers: The Case of the Mortgage Market." *Journal of Finance* 65 (1): 333–60.
- Gertler, Mark, and Peter Karadi. 2011. "A Model of Unconventional Monetary Policy." *Journal of Monetary Economics* 58 (1): 17–34.
- Gertler, Mark, and Nobuhiro Kiyotaki. 2010. "Financial Intermediation and Credit Policy in Business Cycle Analysis." Unpublished.
- Glode, Vincent, Richard C. Green, and Richard Lowery. 2012. "Financial Expertise as an Arms Race." *Journal of Finance* 67 (5): 1723–59.
- Gorton, Gary B., Stefan Lewellen, and Andrew Metrick. 2012. "The Safe-Asset Share." National Bureau of Economic Research Working Paper 17777.
- Gorton, Gary, and Andrew Metrick. 2012. "Securitized Banking and the Run on Repo." *Journal of Financial Economics* 104 (3): 425–51.
- Gorton, Gary, and George Pennacchi. 1990. "Financial Intermediaries and Liquidity Creation." *Journal of Finance* 45 (1): 49–71.
- Gorton, Gary, and Andrew Winton. 2003. "Financial Intermediation." In *Handbook of the Economics of Finance*, Vol. 1A, edited by George M. Constantinides, Milton Harris, and René M. Stulz, 431–552. Amsterdam: Elsevier North Holland.
- Greenwood, Jeremy, Juan M. Sanchez, and Cheng Wang. 2010. "Financing Development: The Role of Information Costs." *American Economic Review* 100 (4): 1875–91.
- Greenwood, Jeremy, Juan M. Sanchez, and Cheng Wang. 2013. "Quantifying the Impact of Financial Development on Economic Development." *Review of Economic Dynamics* 16 (1): 194–215.
- Greenwood, Robin, Samuel G. Hanson, and Jeremy C. Stein. 2011. "A Comparative-Advantage Approach to Government Debt Maturity." Unpublished.
- Greenwood, Robin, and David Scharfstein. 2013. "The Growth of Finance." *Journal of Economic Perspectives* 27 (2): 3–28.
- Guiso, Luigi, Paola Sapienza, and Luigi Zingales. 2004. "The Role of Social Capital in Financial Development." *American Economic Review* 94 (3): 526–56.
- Haldane, Andrew, Simon Brennan, and Vasileios Madouros. 2010. "What Is the Contribution of the Financial Sector: Miracle or Mirage?" In *The Future of Finance: The LSE Report*, edited by Adair Turner et al., 64–104. London: London School of Economics and Political Science.
- Hall, Robert E. 2011. "The High Sensitivity of Economic Activity to Financial Frictions." *Economic Journal* 121 (552): 351–78.
- Hasbrouck, Joel. 2009. "Trading Costs and Returns for US Equities: Estimating Effective Costs from Daily Data." *Journal of Finance* 64 (3): 1445–77.

- He, Zhigu, and Arvind Krishnamurthy.** 2012. "A Model of Capital and Crises." *Review of Economic Studies* 79 (2): 735–77.
- Holmström, Bengt, and Jean Tirole.** 1997. "Financial Intermediation, Loanable Funds, and the Real Sector." *Quarterly Journal of Economics* 112 (3): 663–91.
- Jovanovic, Boyan, and Peter L. Rousseau.** 2001. "Why Wait? A Century of Life before IPO." *American Economic Review* 91 (2): 336–41.
- Jovanovic, Boyan, and Peter L. Rousseau.** 2005. "General Purpose Technologies." In *Handbook of Economic Growth*, Vol. 1B, edited by Philippe Aghion and Steven N. Durlauf, 1181–1224. Amsterdam: Elsevier North Holland.
- Kashyap, Anil K., Raghuram Rajan, and Jeremy C. Stein.** 2002. "Banks as Liquidity Providers: An Explanation for the Coexistence of Lending and Deposit-Taking." *Journal of Finance* 57 (1): 33–73.
- Kiyotaki, Nobuhiro, and John Moore.** 2008. "Liquidity, Business Cycles, and Monetary Policy." Unpublished.
- Krishnamurthy, Arvind.** 2010. "Amplification Mechanisms in Liquidity Crises." *American Economic Journal: Macroeconomics* 2 (3): 1–30.
- Krishnamurthy, Arvind, and Annette Vissing-Jorgensen.** 2012. "The Aggregate Demand for Treasury Debt." *Journal of Political Economy* 120 (2): 233–67.
- Kuznets, Simon.** 1941. *National Income and Its Composition, 1919–1938*. Vol. 1. New York: National Bureau of Economic Research.
- La Porta, Rafael, Florencio Lopez-de-Silanes, Andrei Shleifer, and Robert W. Vishny.** 1998. "Law and Finance." *Journal of Political Economy* 106 (6): 1113–55.
- Lucas, Robert E., Jr.** 2000. "Inflation and Welfare." *Econometrica* 68 (2): 247–74.
- Martin, Robert Fitz.** 1939. *National income in the United States, 1799–1938*. New York: National Industrial Conference Board.
- Mayer, Chris, and Karen Pence.** 2008. "Subprime Mortgages: What, Where, and to Whom?" Federal Reserve Board Staff Working Paper 2008–29.
- Mehra, Rajnish, Facundo Pigullem, and Edward C. Prescott.** 2011. "Costly Financial Intermediation in Neoclassical Growth Theory." *Quantitative Economics* 2 (1): 1–36.
- Mehra, Rajnish, and Edward C. Prescott.** 1985. "The Equity Premium: A Puzzle." *Journal of Monetary Economics* 15 (2): 145–61.
- Merton, Robert C.** 1995. "A Functional Perspective of Financial Intermediation." *Financial Management* 24 (2): 23–41.
- Midrigan, Virgiliu, and Daniel Yi Xu.** 2014. "Finance and Misallocation: Evidence from Plant-Level Data." *American Economic Review* 104 (2): 422–58.
- Modigliani, F., and M. H. Miller.** 1958. "The Cost of Capital, Corporation Finance and the Theory of Investment." *American Economic Review* 48: 261–97.
- Moore, John.** 2011. "Leverage Stacks and the Financial System." Paper presented at the Foundation for the Advancement of Research in Financial Economics conference, Dedham, MA.
- Moore, Kevin B., and Michael G. Palumbo.** 2010. "The Finances of American Households in the Past Three Recessions: Evidence from the Survey of Consumer Finances." Federal Reserve Board Staff Working Paper 2010–06.
- Obstfeld, Maurice, and Alan M. Taylor.** 2002. "Globalization and Capital Markets." National Bureau of Economic Research Working Paper 8846.
- Odean, Terrance.** 1998. "Volume, Volatility, Price, and Profit When All Traders Are above Average." *Journal of Finance* 53 (6): 1887–934.
- Pagnotta, Emiliano, and Thomas Philippon.** 2011. "Competing on Speed." National Bureau of Economic Research Working Paper 17652.
- Philippon, Thomas.** 2008. "The Evolution of the US Financial Industry from 1860 to 2007: Theory and Evidence." Unpublished.
- Philippon, Thomas.** 2012. "Finance versus Wal-Mart: Why are Financial Services so Expensive?" In *Rethinking the Financial Crisis*, edited by Alan S. Blinder, Andrew W. Lo, and Robert M. Solow, 235–46. New York: Russell Sage.
- Philippon, Thomas.** 2015. "Has the US Finance Industry Become Less Efficient? On the Theory and Measurement of Financial Intermediation: Dataset." *American Economic Review*. <http://dx.doi.org/10.1257/aer.20120578>.
- Philippon, Thomas, and Ariell Reshef.** 2012. "Wages and Human Capital in the U.S. Finance Industry: 1909–2006." *Quarterly Journal of Economics* 127 (4): 1551–609.
- Philippon, Thomas, and Ariell Reshef.** 2013. "An International Look at the Growth of Modern Finance." *Journal of Economic Perspectives* 27 (2): 73–96.
- Piketty, Thomas, and Gabriel Zucman.** 2014. "Capital is Back: Wealth-Income Ratios in Rich Countries 1700–2010." *Quarterly Journal of Economics* 129 (3): 1255–1310.

- Pozsar, Zoltan, Tobias Adrian, Adam Ashcraft, and Hayley Boesky.** 2010. "Shadow Banking." Federal Reserve Bank of New York Staff Report 458.
- Rajan, Raghuram G., and Luigi Zingales.** 1998. "Financial Dependence and Growth." *American Economic Review* 88 (3): 559–86.
- Reinhart, Carmen M., and Kenneth S. Rogoff.** 2009. *This Time is Different: Eight Centuries of Financial Folly*. Princeton: Princeton University Press.
- Rhodes-Kropf, Matthew, and David T. Robinson.** 2008. "The Market for Mergers and the Boundaries of the Firm." *Journal of Finance* 63 (3): 1169–211.
- Ritter, Jay R.** 2011. "Initial Public Offerings." Unpublished.
- Scharfstein, David, and Adi Sunderam.** 2011. "The Economics of Housing Finance Reform." In *The Future of Housing Finance: Restructuring the U.S. Residential Mortgage Market*, edited by Martin Neil Baily, 146–198. Washington, DC: Brookings Institution.
- Schularick, Mortiz, and Alan M. Taylor.** 2012. "Credit Booms Gone Bust: Monetary Policy, Leverage Cycles, and Financial Crises, 1870–2008." *American Economic Review* 102 (2): 1029–61.
- Stein, Jeremy C.** 2012. "Monetary Policy as Financial Stability Regulation." *Quarterly Journal of Economics* 127 (1): 57–95.
- Wang, J. Christina, Susanto Basu, and John G. Fernald.** 2009. "A General Equilibrium Asset-Pricing Approach to the Measurement of Nominal and Real Bank Output." In *Price Index Concepts and Measurements*, edited by W. Erwin Diewert, John S. Greenlees, and Charles R. Hulten, 273–320. Chicago: University of Chicago Press.

Has the U.S. Finance Industry Become Less Efficient? On the Theory and Measurement of Financial Intermediation - Data Appendix

Thomas Philippon

A Data Construction

A.1 Income Share of Finance

Figure 4 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 (Carter et al., 2006). 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 Analysis¹.

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 5. It is constructed by using the modern, value-added series from 1955-2012. 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 (Carter et al., 2006). This adjustment is made using the detailed industry data available from (Kuznets, 1941) 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 5 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 7. 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 2012, 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

¹Available online at http://www.bea.gov/industry/gdpbyind_data.htm.

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 (Carter et al., 2006), 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 7.

Two flow variables are shown in Figure 8. The ”Gross Corp. Bond Issuance” variable scales bond issuance of non-financial corporations by GDP. The primary source is (Baker and Wurgler, 2000), who provide data from 1927 to 2008. Historical data from (Carter et al., 2006) is appended from 1910-1926,² and the data for recent years is from the Securities Industry and Financial Markets Association (SIFMA). 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 9. 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 2012. 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 (Carter et al., 2006). 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 9 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 issuances divided by GDP and smoothed using a 3-year moving average are plotted in Figure 13.

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 2012. 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 (Carter et al., 2006) 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 (Carter et al., 2006) 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 10.

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 11. Each of these are an aggregation of six

²Corporate security issues: 1910-1934, series Cj831-837

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 plotted in Figure 7. 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, 2012) and predicting household debt values from 1890 to 1915 by using the relationship 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 2012) 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. 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 be 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 section II C 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 11.

A.6 Intermediation Cost

A time series of the unadjusted intermediation cost of the financial sector is plotted in Figure 3. It is simply the ratio of the domestic income share of the financial sector, plotted on the right-hand panel of Figure 5, over the composite output measure of the financial sector, combining the flow and level output measures plotted in Figure 11 and as described above.

A.7 Quality Adjustments

Section III of the paper as well as section 2 of Online Appendix discuss the possibility that the financial sector may face increasing marginal costs in supplying credit to riskier firms or households. The adjustments are displayed on Figures 14 and 15 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). The (Piketty and Saez, 2003) time series is available online³ 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 16. 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 17.

³from Saez's homepage <http://elsa.berkeley.edu/~saez/> as of November 2011.

References

- Baker, Malcolm, and Jeffrey Wurgler.** 2000. “The Equity Share in New Issues and Aggregate Stock Returns.” *Journal of Finance*, 55: 2219–2257.
- Carter, Susan, Scott Gartner, Michael Haines, Alan Olmstead, Richard Sutch, and Gavin Wright,** ed. 2006. “Historical Statistics of the United States Millennial Edition Online.” Cambridge University Press.
- Jovanovic, Boyan, and Peter L. Rousseau.** 2005. “General Purpose Technologies.” Elsevier. Chapter 18.
- Kuznets, Simon.** 1941. “National Income and Its Composition, 1919-1938.” National Bureau of Economic Research.
- Piketty, Thomas, and Emmanuel Saez.** 2003. “Income Inequality in the United States, 1913-1998.” *Quarterly Journal of Economics*, 118(1): 1–39.
- Schularick, Moritz, and Alan M Taylor.** 2012. “Credit Booms Gone Bust.” *American Economic Review*, 102(2): 1029–61.

Has the U.S. Finance Industry Become Less Efficient? On the Theory and Measurement of Financial Intermediation - Online Appendix

Thomas Philippon

1 Benchmark Model

In this section I briefly describe a neoclassical growth model with financial intermediation.¹ This model clarifies the comparative statics and provides guidance for the aggregation of financial claims. The model economy consists of households, a non-financial business sector, and a financial intermediation sector. Long term growth is driven by labor-augmenting technological progress $A_t = (1 + \gamma)A_{t-1}$. In the benchmark model borrowers are homogenous, which allows a simple characterization of equilibrium intermediation. I discuss heterogeneity and quality adjustments later.

1.1 Informal Discussion of the Benchmark Model

To organize the discussion I use a simple model economy consisting of households, a non-financial business sector, and a financial intermediation sector. The model assumes homogenous borrowers within each group. Heterogeneity and quality adjustments are discussed in details in Section 2. This section simply highlights some important properties of the model.

Income of the finance industry. The model assumes that financial services are produced under constant returns to scale. The income of the finance industry y_t^f is given by

$$y_t^f = \psi_{c,t}b_{c,t} + \psi_{m,t}m_t + \psi_{k,t}k_t, \quad (1)$$

where $b_{c,t}$ is consumer credit outstanding, m_t are holdings of liquid assets, and k_t is the value of intermediated corporate assets. The parameters $\psi_{i,t}$'s are the unit cost of intermediation, pinned down by the intermediation technology. The model therefore says that the income of the finance industry is proportional to the quantity of intermediated assets, properly defined. The model predicts no income effect, i.e., no tendency for the finance income share to grow with per-capita GDP. This does not mean that the finance income share should be constant, since the ratio of assets to GDP can change. But it says that the income share does not grow mechanically with total factor productivity. This is consistent with the evidence presented below.

²

Corporate finance. As far as corporate finance is concerned, the model is fundamentally a user cost model. The key equation is

$$k_t^\alpha = \frac{1 - \alpha}{r_t + \delta + \psi_{k,t}}, \quad (2)$$

¹It is critical to model financial services explicitly. It is well known that the properties of two-sectors models depend on the elasticity of substitution between the two sectors (Baumol, 1967). For instance, the nominal GDP share of sector i increases with relative technological progress in sector i if and only if the elasticity of substitution is less than one. In the context of financial intermediation, I will show that the elasticity depends both on the shape of the distribution of borrowers and on the efficiency of the supply of financial services. It is therefore not possible to take this elasticity as an (exogenous) parameter.

²The fact that the finance share of GDP plotted in figure 4 of the article is the same in 1925 and in 1980 makes it already clear that there is no mechanical relationship between GDP per capita and the finance income share. Similarly, Bickenbach et al. (2009) show that the income share of finance has remained remarkably constant in Germany over the past 30 years. More precisely, using KLEMS for Europe (see O'Mahony and Timmer (2009)) one can see that the finance share in Germany was 4.3 percent in 1980, 4.68 in 1990, 4.19 in 2000, and 4.47 percent in 2006.

where α is the labor share, δ the depreciation rate, and r the equilibrium interest rate. Improvements in corporate finance (a decrease in ψ_k) lead to a lower user cost of capital, a higher capital/output ratio, and a higher real wage. Despite its simplicity, this framework can account for business risk management. One well-understood benefit of risk management is to reduce the cost of financial distress, which is the net present value of the deadweight losses incurred in states of the world where firms are in financial distress. These deadweight losses include bankruptcy costs and foregone investment opportunities (see Froot, Scharfstein and Stein (1993) for a classic contribution, and Almeida and Philippon (2007) for a discussion of how to capitalize these costs). Standard models assume that financial distress can destroy a fraction of the firm's assets. From an ex-ante perspective, this is equivalent to a higher depreciation rate δ . Conversely, risk management offers one way to lower δ . This is important because there is often some confusion about how to take into account the economic value of financial derivatives. The value of risk management is capitalized in the market value of assets and the user cost $r + \delta + \psi_k$ should account properly for improvements in risk management.

Household finance. A significant part of the growth of the finance industry over the past 30 years is linked to household credit. The model provides a simple way to model household finance. In the model, borrowing costs act as a tax on future income. If ψ_c is too high, no borrowing takes place and the consumer credit market collapses. The model also incorporates liquidity services provided by specific liabilities (deposits, checking accounts) issued by financial intermediaries.

1.2 Technology and Preferences

The model economy consists of households, a non-financial business sector, and a financial intermediation sector. In the model, the finance industry provides three types of services to households and firms: liquidity, credit, and asset management. Households hold the corporate capital stock via intermediaries. In addition, households borrow and lend from each other. Since household debt has an important life-cycle component (i.e., mortgages), I consider a setup with two types of households: some households are infinitely lived, the others belong to an overlapping generations structure.³

Households in the model do not lend directly to one another. They lend to intermediaries, and intermediaries lend to firms and to other households.

Long-Lived Households

Long-lived households (index l) own the capital stock and have no labor endowment. Liquidity services are modeled with money in the utility function (using a cash-in-advance framework gives similar results). The households choose consumption C and holdings of liquid assets M to maximize

$$\mathbb{E} \sum_{t \geq 0} \beta^t u(C_t, M_t).$$

I specify the utility function as $u(C_t, M_t) = \frac{(C_t M_t^\nu)^{1-\rho}}{1-\rho}$. As argued by Lucas (2000), these homothetic preferences are consistent with the absence of trend in the ratio of real balances to income in U.S. data, and the constant relative risk aversion form is consistent with balanced growth. Let r be the interest rate received by savers. The budget constraint becomes

$$S_t + C_t + \psi_{m,t} M_t \leq (1 + r_t) S_{t-1},$$

where ψ_m is the price of liquidity services, and S are total savings.⁴ The Euler equation of long lived households $u_C(t) = \beta E_t [(1 + r_{t+1}) u_C(t+1)]$ can then be written as

$$M_{l,t}^{\nu(1-\rho)} C_{l,t}^{-\rho} = \beta E_t \left[(1 + r_{t+1}) M_{l,t+1}^{\nu(1-\rho)} C_{l,t+1}^{-\rho} \right].$$

³The pure infinite horizon model and the pure OLG model are both inadequate. The infinite horizon model misses the importance of life-cycle borrowing and lending. The OLG model ignores bequests, and in the simple two-periods version households do not actually borrow: the young ones save, and the old ones eat their savings. The simplest way to capture all these relevant features is the mixed model. The standard interpretation is that long-lived households have bequest motives, and are therefore equivalent to infinitely lived agents. See also Mehra, Piguillem and Prescott (2011) for a model where households save for retirement over an uncertain lifetime.

⁴See Lucas and Stokey (1987) and Sargent and Smith (2010) for a discussion of cash-in-advance models. Lucas (2000) uses the framework of Sidrauski (1967) with a more flexible functional form of the type $\left(C_t \varphi \left(\frac{M_t}{C} \right) \right)^{1-\rho}$. I use a Cobb-Douglas aggregator for simplicity given the complexity of the rest of the model. A more important difference with the classical literature on money demand is that I do not focus on inflation. Households save S at a gross return of $1 + r$, while liquid assets yield $(1 + r)/(1 + \psi_m)$. So this model implies a constant spread between the lending rate and the rate on liquid assets. This is consistent with my interpretation of liquidity as not only money, but also money market funds shares and repurchase agreements.

The liquidity demand equation $u_M(t) = \psi_{m,t} u_C(t)$ is simply

$$\psi_{m,t} M_{l,t} = \nu C_{l,t}.$$

Overlapping Generations

The other households live for two periods and are part of an overlapping generation structure. The young (index 1) have a labor endowment η_1 and the old (index 2) have a labor endowment η_2 . We normalize the labor supply to one: $\eta_1 + \eta_2 = 1$. The life-time utility of a young household is $u(C_{1,t}, M_{1,t}) + \beta u(C_{2,t+1}, M_{2,t+1})$. I consider the case where they want to borrow when they are young (i.e., η_1 is small enough). In the first period, its budget constraint is $C_{1,t} + \psi_{m,t} M_{1,t} = \eta_1 W_{1,t} + (1 - \psi_{c,t}) B_t^c$. The screening and monitoring cost is $\psi_{c,t}$ per unit of borrowing. In the second period, the household consumes $C_{2,t+1} + \psi_{m,t+1} M_{2,t+1} = \eta_2 W_{t+1} - (1 + r_{t+1}) B_t^c$. The Euler equation for OLG households is

$$(1 - \psi_{c,t}) M_{1,t}^{\nu(1-\rho)} C_{1,t}^{-\rho} = \beta \mathbb{E}_t \left[(1 + r_{t+1}) M_{2,t+1}^{\nu(1-\rho)} C_{2,t+1}^{-\rho} \right].$$

Their liquidity demand is identical to the one of long-lived households.

Non Financial Businesses

Non-financial output is produced under constant returns technology, and for simplicity I assume that the production function is Cobb-Douglas:⁵

$$F(A_t n_t, K_t) = (A_t n_t)^\alpha K_t^{1-\alpha}.$$

The capital stock K_t depreciates at rate δ , is owned by the households, and must be intermediated. Let $\psi_{k,t}$ be the unit price of corporate financial intermediation. Section 2 derives the demand for intermediation services from a standard moral hazard model, but for now I take $\psi_{k,t}$ as a parameter. Non financial firms therefore solve the following program $\max_{n,K} F(A_t n, K) - (r_t + \delta + \psi_{k,t}) K - W_t n$. Capital demand equates the marginal product of capital to its user cost:

$$(1 - \alpha) \left(\frac{A_t n_t}{K_t} \right)^\alpha = r_t + \delta + \psi_{k,t}. \quad (3)$$

Similarly, labor demand equates the marginal product of labor to the real wage:

$$\alpha \left(\frac{A_t n_t}{K_t} \right)^{\alpha-1} = \frac{W_t}{A_t}. \quad (4)$$

Financial Intermediation

Philippon (2012) discusses in details the implications of various production functions for financial services. When financial intermediaries explicitly hire capital and labor there is a feed-back from intermediation demand onto the real wage. This issue is not central here, and I therefore assume that financial services are produced from final goods with constant marginal costs. The income share of financial intermediaries is then

$$\phi_t = \psi_{c,t} \frac{B_{c,t}}{Y_t} + \psi_{m,t} \frac{M_t}{Y_t} + \psi_{k,t} \frac{K_t}{Y_t}$$

where Y_t is aggregate GDP and $B_{c,t}$, M_t and K_t have been described above.

1.3 Equilibrium Comparative Statics

An *equilibrium* in this economy is a sequence for the various prices and quantities listed above such that households choose optimal levels of credit and liquidity, financial and non financial firms maximize profits, and the labor and capital markets clear. This implies $n_t = 1$ and

$$S_t = K_{t+1} + B_t^c.$$

⁵Philippon (2012) discusses the consequences of assuming a different production function for the industrial sector. The key parameter is the elasticity of substitution between capital and labor, which is 1 under Cobb-Douglas technology. Qualitatively different results only happen for elasticity values above 6, which is far above the range of empirical estimates. Thus assuming a Cobb-Douglas technology does not entail much loss of generality.

Let us now characterize an equilibrium with constant productivity growth in the non-financial sector (γ) and constant efficiency of intermediation (ψ). On the balanced growth path, M grows at the same rate as C . The Euler equation for long-lived households becomes $1 = \beta \mathbb{E}_t \left[(1 + r_{t+1}) \left(\frac{C_{t+1}}{C_t} \right)^{\nu(1-\rho)-\rho} \right]$, so the equilibrium interest rate is simply pinned down by

$$\beta(1+r) = (1+\gamma)^\theta. \quad (5)$$

where $\theta \equiv \rho - \nu(1 - \rho)$. Let lower-case letters denote de-trended variables, i.e. variables scaled by the current level of technology: for capital $k \equiv \frac{K_t}{A_t}$, for consumption of agent i $c_i \equiv \frac{C_{i,t}}{A_t}$, and for the productivity adjusted wage $w \equiv W_t/A_t$. Since $n = 1$ in equilibrium, equation (3) becomes

$$k^\alpha = \frac{1 - \alpha}{r + \delta + \psi_k}.$$

Non financial GDP is $y = k^{1-\alpha}$, and the real wage is

$$w = \alpha k^{1-\alpha} = \alpha y.$$

Given the interest rate in (5), the Euler equation of short lived households is simply

$$c_1 = (1 - \psi_c)^{\frac{1}{\theta}} c_2. \quad (6)$$

If ψ_c is 0, we have perfect consumption smoothing: $c_1 = c_2$ (remember these are de-trended consumptions). In addition, all agents have the same money demand $\psi_m m_i = \nu c_i$. The budget constraints are therefore $(1 + \nu) c_1 = \eta_1 w + (1 - \psi_c) b$ and $(1 + \nu) c_2 = \eta_2 w - \frac{1+r}{1+\gamma} b$. We can then use the Euler equations and budget constraints to compute the borrowing of young households

$$\frac{b_c}{w} = \frac{(1 - \psi_c)^{\frac{1}{\theta}} \eta_2 - \eta_1}{1 - \psi_c + (1 - \psi_c)^{\frac{1}{\theta}} \frac{1+r}{1+\gamma}}. \quad (7)$$

Borrowing costs act as a tax on future labor income. If ψ_c is too high, no borrowing takes place and the consumer credit market collapses. Household borrowing increases with the difference between current and future income, captured by $\eta_2 - \eta_1$. Liquidity demand is

$$m = \frac{\nu c}{\psi_m}.$$

and aggregate consumption is

$$c = \frac{1}{1 + \nu} (w - \psi_c b_c + (r - \gamma) k). \quad (8)$$

The following proposition summarizes the predictions of the theory.

Proposition 1. *The equilibrium has the following features*

(i) *Along a balanced growth path with constant intermediation technology, constant demographics, and constant firms' characteristics, the finance share of GDP and the financial ratios m/y , b^c/y and k/y are constant.*

(ii) *Improvements in corporate finance increase y , w , k/y , c/y and m/y , but leave b^c/y constant;*

(iii) *Improvements in household finance increase b^c/y , c/y and m/y , but do not affect k ;*

(iv) *Increases in the demand for intermediation increase the finance income share ϕ while supply shifts have an ambiguous impact.*

Proof. Point (i): it is clear from the equations of the model. Point (ii): when ψ_k goes down, k , y and w go up, k/y goes up because of decreasing returns to capital, but equation (7) shows that b_c/y remain constant since $w = \alpha y$. From (8), c/y increases because k/y increases, and so does m/y . Point (iii): when ψ_c goes down, b_c , c and m go up. Since the user cost of capital is not affected, k and w are constant. Point (iv): suppose η_2 goes up, and $\eta_1 = 1 - \eta_2$ goes down. From (7), b_c goes up while k , w and y are unchanged. Then c and m go down, but one can check that $\psi_c b_c + \psi_m m$ goes up and therefore ϕ increase. On the other hand, suppose now that ψ goes down. There are two effects: financial services are cheaper which pushes y^f down, but more services will be provided in equilibrium. So the net impact on y^f is theoretically ambiguous. \square

That homogeneity in production is required for balanced growth is not surprising. What is more interesting is that it is sufficient even if the production technologies differ between the financial and non-financial sectors (see Philippon (2012) for detailed discussions).

1.4 Risk management

The market for financial derivatives is extremely large. Since these contracts are in zero net supply, however, they do not enter *directly* into my calculations of intermediated assets. The question is: Should they? The answer is essentially no, because the benefits of derivatives are already *indirectly* taken into account.

One thing is clear: it would make no sense to count derivatives at face value. Rather, one should take the perspective of standard economic theory and recognize that derivatives can add real value in one of two ways: (i) risk sharing; (ii) price discovery. Bai, Philippon and Savov (2011) discuss price discovery so let me focus on how risk management affects my measures. Risk management among banks lowers intermediation costs as banks lay off excess risk inventories when making markets. This type of risk sharing among financial intermediaries does not create any bias in my measurements. Improvements in risk management would simply lead to lower borrowing costs and cheaper financing, and this would be correctly captured by the model.⁶

Risk management by the non-financial sector is more subtle, but the user cost framework is still the right approach. To see why, decompose the depreciation rate into an exogenous physical depreciation δ_p and an expected cost of distress δ_f . We can then endogenize the expected cost of distress by assuming that the technological production frontier of the finance industry is $\delta_f^2 + \psi^2 \geq \chi^2$, where ψ is the cost of risk management and χ indexes the production productivity frontier of risk management services. This captures the idea that it is possible to decrease financial distress by spending more on risk management. The optimal choice is simply to minimize the user cost of capital, i.e.,

$$\min \delta_f + \psi \text{ s.t. } \delta_f^2 + \psi^2 \geq \chi^2.$$

This then leads to $\delta_f = \psi = \chi/\sqrt{2}$. Then the total financial cost is $\delta_f + \psi = \chi\sqrt{2}$, and in steady state we have

$$k = \left(\frac{1 - \alpha}{r + \delta_p + \chi\sqrt{2}} \right)^{\frac{1}{\alpha}}.$$

Improvements in risk management increase k , just like generic improvements in intermediation. The finance industry earns $\psi k = \frac{\chi}{\sqrt{2}}k$, so my measure of unit cost would simply recover the evolution of χ .

In the neoclassical growth model with homogenous firms, there is no value of the firm as a going concern. In the next section, I consider an extension with heterogenous firms and decreasing returns at the firm level, where the value of some firms exceeds the replacement cost of their capital. In general we can think of firm value (per unit of capital) as

$$V = \pi + \frac{1 - \delta_p - \delta_f - \psi}{1 + r} V'$$

where V is the current value, V' the future value, π is the profit rate, δ_f the cost of financial distress and ψ the cost of risk management. Using the same intermediation technology as above, the optimal risk management yields $\delta_f = \psi = \chi/\sqrt{2}$. Solving for V in steady state, we have

$$V = \frac{(1 + r)\pi}{r + \delta_p + \chi\sqrt{2}}$$

The important point here is that improvement in risk management would be capitalized in the market value of the firm. This provides a clear argument for using the market value of equity in the calculations. A caveat here is that I treat risk management and price discovery as two separate issues, but they need not be. In DeMarzo and Duffie (1991) for instance, financial hedging is fundamentally linked to private information about firm value, and in DeMarzo and Duffie (1995) hedging interacts with incentives and accounting disclosure. Those complex and fascinating issues are beyond the scope of this paper.

⁶For instance, consider the following example. Without derivatives, corporation A borrows from bank B and bank B retains the credit and duration risks on its books. With derivatives, bank B buys insurance against credit risk from fund C using a CDS. The sum of B and C holds exactly the same risk. Absent other frictions, the two models are equivalent. Now suppose there are frictions that rationalize why B and C should be separate entities, and why they gain from trading with each other (i.e., B has a comparative advantage at managing duration risk, and C at managing credit risk). Then the existence of CDS contracts can improve risk sharing among intermediaries, lower the risk premia, and lead to a decrease in the borrowing costs of A. Hence, with free entry, the total income going to intermediaries B+C would decrease. This could then increase the demand for borrowing, as explained earlier. All these effects would be captured by the model: either borrowing costs would go down, or borrowing volumes would go up, or both. In all cases, my approach would register an increase in efficiency.

1.5 Calibration of basic model

If we could observe the income flows $y_{i,t}^f$, associated with the three fundamental sources of revenues $i = b, m, k$, we would simply compute the unit cost as, for instance: $\psi_{c,t} = \frac{y_{c,t}^f}{b_{c,t}}$ where $y_{c,t}^f$ would be the income generated by credit intermediation for consumers. Unfortunately, there is no satisfactory way to link a particular income to a particular activity, especially over long periods of time.⁷ This precludes a direct estimation of the $\psi_{i,t}$'s. We only observe the total income of the finance industry, y_t^f described in Section 1 of the article.

To make progress therefore requires new assumptions. In this Section, I assume that any change in efficiency of financial intermediation is shared proportionately among the three main activities:

Assumption CRC (constant relative costs) : $\psi_{i,t} = \mu_i \psi_t$, with the normalization $\mu_c = 1$.

This assumption means that the *relative* difficulty of making a corporate loan versus a household loan remains constant over time, even though the unit costs can change. Assumption CRC will be relaxed in Section ??, albeit at the cost of much greater complexity. But for now, it allows me to defined the (weighted) quantity of intermediated assets as

$$q_t \equiv b_{c,t} + \mu_m m_t + \mu_k k_t.$$

Calibrating the relative costs. What remains to be done is to estimate the weights μ_m and μ_k . We can obtain indirect estimates by using the first order conditions of the model together with microeconomic evidence on the prices of various financial services. The benchmark interest rate in the economy is r . Liquid assets yield $r - \psi_m$ and (consumer) loan rates are equal to $r + \psi_c$. Direct evidence suggests that the reference rate is roughly halfway between the deposit rate and the loan rate (net of expected default). This implies $\mu_m = \mu_c$ and, given my estimated unit costs, that ψ_m is between 1.5 and 2 percent, which is consistent with various data sources.⁸

Corporate intermediation is made of equity and debt financing, hence $\mu_k k = \mu_k^b b_k + \mu_k^e e_k$, where e_k is corporate equity and b_k is corporate debt. Altinkilic and Hansen (2000) report fees of 3 to 4 percent for equity issuances and of 1 percent for bond issuances.⁹

For flows of new issuances, I therefore assume $\mu_k^e = 3.5\mu_k^b$. These numbers seem to be in line with recent reports by large investment banks. For instance, JP Morgan's 2010 annual report suggests underwriting fees around 0.70 percent for debt, and around 2.46 percent for equity (see below). Finally, as a benchmark I set $\mu_k^b = \mu_c$ so that it is equally difficult to extend credit to firms and to households.

JP Morgan 2010 According to its 2010 annual, total net revenue for JPM Co was \$103 billion, \$51b of interest income and \$52b of non-interest income. The investment bank earned \$26 billion, 15 from fixed income markets, 5 from equity markets, and a bit more than 6 in fees. Of the \$26b, non interest income accounted for 18, including 6.2b

⁷There is an empirical problem and a conceptual problem. Empirically, our data is organized by industry (e.g., Securities, Credit Intermediation), not by function and even less by end-user. Even obtaining detailed measures of gross output is challenging. See Greenwood and Scharfstein (2013) for an enlightening discussion. But this is not only an issue of accounting. Even if we had all the data imaginable, we would still need to decide how to allocate costs among many shared activities: hedging and risk management, trading, over-head labor, etc. And financial tasks are deeply intertwined. Insurance companies and pension funds perform their own independent credit analysis. Banks act as market makers. Investment banks behave as hedge funds. In addition, the mapping from industry to tasks has changed over time with the development of the originate and distribute model in banking. Therefore the problem runs even deeper if we want to make long run comparisons.

⁸The "user cost and reference rate" approach is the one used by the Bureau of Labor Statistics (BLS) to construct Producer Price indexes for Commercial Banking and Savings Institutions. The services for which indexes are available include: Loans, Deposits, Trust services, and Other banking services. To measure prices in these industries, PPI has implemented a user-cost methodology. The user cost for a financial service is the difference between the revenue it generates and the sum of its implicit and explicit costs. To measure these costs, interest is allocated between loans and deposits by means of a reference rate. Hood (2013) shows that, in 2008Q1, the loan rate (net of expected default) is 6.18 percent, the deposit rate is 2.74, and the reference rate is 4.35 percent. This would be consistent with $r = 0.0435$, $\psi_c = 0.0183$ and $\psi_m = 0.0161$. In figure 6.4 of Fixler (2009) the reference rate is also halfway between the deposit rate and the loan rate. Table 1 presents other relevant rates and returns. Over the period 2002-2011, the Vanguard Short Term Treasury fund has returned 3.65 percent, with an expense ratio of 0.22 (Vanguard data accessed on March 11, 2012). This is the opportunity cost of cash, but cash is a relatively small fraction of liquid assets. Over the same period, the Vanguard Prime Money Market fund has returned 1.82 percent with an expense ratio of 0.20. The difference in returns between these two essentially risk free instruments is 1.8 percent. Similarly, the difference between the 1-month CD rate and the 10-year Bond rate is 1.8 percent, while the difference between the 10-year Bond and the conventional mortgage is 1.76 percent. One can also interpret ψ_m as the cost of creating liquid assets. This cost can be charged as a redemption fee for investors. This fee is around 2 percent and is consistent with the trading costs incurred by mutual funds upon withdrawals (see Chen, Goldstein and Jiang (2010) for a discussion).

⁹For households we can look at the mortgage market. Sirmans and Benjamin (1990) report fees of 0.50 to 0.70 percent. For other types of consumer credit these fees are certainly larger.

Table 1: U.S. Interest Rates and Returns, 2002-2011

Variable	Value (percent)
Average 3M Treasury Bills	1.79
Average 1M Certificate of Deposits	2.14
Average 1Y Gov. Bond	2.10
Average 10Y Gov. Bond	3.95
Average Aaa Corporate	5.47
Average Prime Bank Loan	5.02
Average 30Y Conventional Mortgage	5.71
Average Baa Corporate	6.64
Vanguard Prime Money Market Fund Return	1.82
Vanguard Short Term Treasury Fund Return	3.65

Source: FRED, and Vanguard. All values are average over 2002-2011.

in fees (3.1 and 1.6 for debt and equity underwriting, and 1.5 for advisory fees), 8.4b from principal transactions, and 2.5b from asset management fees. For its private clients, the investment bank raised \$440b in debt and \$65b in equity. This suggests underwriting fees of $3.1/440 = 0.70$ percent for debt, and $1.6/65 = 2.46$ percent for equity. The cost of equity underwriting is therefore about 3.5 times the cost of debt underwriting. The bank also raised \$90b for governments and non-profits. The bank advised 311 announced M&A (a 16 percent market share). The bank also loaned or arranged \$350b.

2 Quality Adjustments

This section presents the model with heterogenous firms and households. It extends the model of Appendix A.

2.1 Firms

This section describes what happens within a period, so I suppress the time index (remember that in the OLG setup, one period represents many years). There are k firms. Firm i is endowed with $x_i A$ and needs to borrow $(1 - x_i)A$ to operate a technology that produces according to $f(n) = An^\alpha$. With Cobb-Douglas technology, we get net income

$$\pi(w) = (1 - \alpha) \left(\frac{\alpha}{w} \right)^{\frac{1}{1-\alpha}}$$

and labor demand $n = \left(\frac{\alpha}{w} \right)^{\frac{1}{1-\alpha}}$.

There are two frictions in capital markets. The first is a proportional cost φ as in Appendix A, which capture holding costs and asset management fees. The second friction comes from moral hazard. Firms' owners can divert resources and monitoring by intermediaries is used to reduce the risk of diversion. If the firm behaves well, it pays back its outside investors $(1 + r)(1 - x)$ and insiders receive

$$\pi(w) + 1 - \delta - \varphi - (1 - x)(1 + r) - \zeta\mu = \pi(w) - \delta - \varphi - r - \zeta\mu_k + (1 + r)x,$$

where ζ is the unit cost of monitoring and μ the quantity of monitoring used by the firm. If the firm cheats, outside investors receive nothing and inside investors keep $(1 + r)\xi - \zeta\mu - \mu$, where ξ measures the degree of diversion.¹⁰ The incentive constraint is therefore $\pi(w) - \delta - r + (1 + r)x \geq (1 + r)\xi - \mu$. The firm seeks to minimize the cost of monitoring subject to a break-even constraint and an incentive constraint:

$$\begin{aligned} \min_{\mu \geq 0} \zeta\mu & \quad s.t. \\ \pi(w) & \geq r + \delta + \varphi + \zeta\mu, \\ \pi(w) & \geq r + \delta + \varphi + (1 + r)(\xi - x) - \mu. \end{aligned}$$

¹⁰This formula assumes that firms cannot divert bankers' fees. The analysis is essentially the same if they can, one must simply carry an extra term $\zeta\mu_k$ in the formulas.

The first constraint binds for marginal firms, i.e., firms that are indifferent between entering and staying out, and all firms must satisfy their incentive constraints. This defines the required level of monitoring as a function of aggregate parameters (r, π) and of the firm's level of free cash flows x :

$$\mu(x) = \max\langle 0; r + \delta + \varphi - \pi(w) + (1+r)(\xi - x) \rangle. \quad (9)$$

We consider a model with two types of firms, $i = l, h$ with $x_l < x_h$. We study equilibria where the number of high cash ventures is (exogenously) given by k_h and where l is the marginal type (i.e., k_h is not too large enough to exhaust the investment needs of the economy). There is free entry of low types, therefore we have

$$\pi(w) = r + \delta + \varphi + \zeta\mu(x_l),$$

and substituting in the IC constraint we get

$$\mu(x_l) = \frac{1+r}{1+\zeta}(\xi - x_l).$$

This pins down the required profit rate $\pi(w)$, and therefore the equilibrium wage:

$$\pi(w) = r + \delta + \varphi + \frac{\zeta}{1+\zeta}(1+r)(\xi - x_l).$$

The h -firms earn rents since $\mu(x_h) < \mu(x_l)$. From (9) and assuming that $\mu_h > 0$, the difference in monitoring requirements is

$$\mu_l - \mu_h = (1+r)(x_h - x_l).$$

The quantity of monitoring produced by financial intermediaries is given by

$$\bar{\mu} = \sum_{j=l,h} k_j \mu_j.$$

The income of the finance industry is made of direct intermediation income (monitoring, screening) and asset management fees

$$y_{k,t}^f = \varphi_t k_t + \zeta_t \bar{\mu}_t.$$

Focusing on aggregate borrowing, we have $\bar{b} = k_h(1 - x_h) + k_l(1 - x_l)$, so the intensity of monitoring, $\bar{\mu}/\bar{b}$ is

$$\frac{\bar{\mu}}{\bar{b}}(s) = \frac{\mu_h + (1+r)(x_h - x_l)s}{1 - x^h + (x_h - x_l)s}$$

where $s \equiv \frac{k_l}{k_l + k_h}$ is the fraction of high monitoring (low-cash) firms in aggregate investment. The ratio of income over intermediated assets is for the monitoring activity is

$$\zeta_t \frac{\bar{\mu}_t}{\bar{b}_t}$$

The parameter of interest is ζ_t which captures the true efficiency of financial intermediation. In the data I measure $\frac{\zeta_t \bar{\mu}_t}{\bar{b}_t}$. To recover ζ_t I therefore need to estimate $\frac{\bar{\mu}}{\bar{b}}(s_t)$.

2.2 Households

I consider the case of log preferences $\rho = 1$. Long-lived households are the same as before, therefore

$$\beta(1+r) = 1 + \gamma$$

Short-lived households have a desire to borrow. There is a fixed cost to borrowing, in addition to the marginal cost φ . Households differ in their incomes. As before, they value their stream of consumptions as

$$V_t = \log C_{1,t} + \nu \log M_{1,t} + \beta \log C_{2,t+1} + \beta \nu \log M_{2,t+1}.$$

Money demand is unchanged, so we still have

$$\psi_{m,t} M_{i,t} = \nu C_{i,t}.$$

Normalizing by aggregate productivity A_t , we see that

$$V_t = \log c_{1,t} + \beta \log c_{2,t+1} + \Xi_t,$$

where $\Xi_t \equiv (1 + \nu)(\log A_t + \beta \log A_{t+1}) + \nu \log \frac{\nu}{\psi_{m,t}} + \nu \beta \log \frac{\nu}{\psi_{m,t+1}}$ only depends on aggregate quantities. The households' choices therefore really depend upon $\log(c_1) + \beta \log(c_2)$, assuming as before constant TFP growth.

Autarky If the household stays out of the credit market, its budget constraints are simply $(1 + \nu) c_1 = \eta_1 w$ and $(1 + \nu) c_2 = \eta_2 w$. The value of autarky is therefore

$$V_{aut} = \log(\eta_1) + \beta \log(\eta_2) + (1 + \beta) \log \frac{w}{1 + \nu}$$

Credit If the household enters the credit market, it pays the fixed cost κw at time 2. The budget constraints are $(1 + \nu) c_1 = \eta_1 w + (1 - \varphi) b$ and $(1 + \nu) c_2 = \eta_2 w - \kappa w - \frac{1+r}{1+\gamma} b$, and its Euler equation is therefore

$$c_1 = (1 - \varphi) c_2.$$

Obviously, if $\psi = 0$, we have $c_1 = c_2$. From these equation we can compute the amount borrowed as

$$\frac{b_c}{w} = \frac{\eta_2 - \kappa - \frac{\eta_1}{1-\varphi}}{1 + \frac{1+r}{1+\gamma}}. \quad (10)$$

Clearly, borrowing requires $\eta_2 > \kappa + \frac{\eta_1}{1-\varphi}$. The maximum amount of borrowing, if $\varphi = \kappa = 0$ is $\frac{b_c}{w} = \frac{\eta_2 - \eta_1}{1 + \frac{1+r}{1+\gamma}}$. The value of entering the credit market is

$$V_{bor} = (1 + \beta) \log(c_1) - \beta \log(1 - \varphi).$$

We can compute the consumption levels from the budget constraints:

$$c_1 = \frac{w}{1 + \nu} \frac{(1 + r) \eta_1 + (1 - \varphi) (1 + \gamma) (\eta_2 - \kappa)}{2 + r + \gamma}.$$

Households' choices I parameterize the model in the following way. Let η be a random variable with mean of $\frac{1}{1+\lambda}$ drawn at birth. $\eta_1 = \eta$ and $\eta_2 = \lambda \eta$ with $\lambda > 1$. The choice of entering the credit market or not depends on

$$\Delta \equiv V_{bor} - V_{aut}$$

Substituting the above expressions, we get

$$\Delta = (1 + \beta) \log \left(\frac{(1 + r + (1 - \varphi) (1 + \gamma) \left(\lambda - \frac{\kappa}{\eta} \right))}{2 + r + \gamma} \right) - \beta \log \lambda (1 - \varphi)$$

We see that

$$\Delta > 0 \iff \eta > \hat{\eta}$$

where the cutoff $\hat{\eta}$ solves

$$\lambda - \frac{\kappa}{\hat{\eta}} = \frac{(2 + r + \gamma) (\lambda (1 - \varphi))^{\frac{\beta}{1+\beta}} - (1 + r)}{(1 - \varphi) (1 + \gamma)}.$$

Aggregate outcome Liquidity demand is

$$m = \frac{\nu c}{\psi_m}.$$

and aggregate consumption is

$$c = \frac{1}{1 + \nu} (w - \varphi \bar{b}_c - \kappa w (1 - F(\hat{\eta})) + (r - \gamma) k).$$

where aggregate household debt is

$$\bar{b}_c = \frac{1 + \gamma}{2 + r + \gamma} \int_{\eta > \hat{\eta}} \left(\left(\lambda - (1 - \varphi)^{-1} \right) \eta - \kappa \right) dF(\eta).$$

Finance income from consumer credit is

$$\varphi \bar{b}_c + \kappa w (1 - F(\hat{\eta})).$$

2.3 Calibration and Mapping from the Model to the Data

Table 4 presents the parameters. Some parameters are set using values that are standard in the literature. The finance-specific parameters are chosen to match micro-data and a set of moments. The moments are presented in the second row of the table, and the implied parameters in the last row. I calibrate the model using 1989 as a reference year because the Survey of Consumer Finance is available for that year.

Table 2: Standard Parameters

Rate	Deprec.	Growth	Labor Sh.	CRRA
$r = 0.05$	$\delta = 0.1$	$\gamma = 0.02$	$\alpha = 0.7$	$\rho = 1$

Corporate Finance In the model presented here, there are two types of intermediation. Internal funds cover a fraction x_i of the capital needs, with $x_h > x_l$. The remaining $1 - x_i$ is external, and I refer to it as external finance. The management fee φ applies to the entire stock of capital while the monitoring cost applies only to $b_k = 1 - x$. To map these quantities with the data, I make the traditional assumption that external finance is debt and new issuances of debt and equity, while outstanding equity is internal. I therefore define external finance as total business intermediation minus the value of existing equity, or equivalently as the stock of debt plus issuances of new debt and equity. In 1989, the stock of existing firm debt is 0.665 GDP. Taking into account the flows of new debt and equity, monitored finance is 0.806 GDP. This is the target value for \bar{b}_k/y in the model.

Following Greenwood and Scharfstein (2013), I set the asset management fee to $\varphi = 1$ percent and I apply this fee with the appropriate accounting corrections.¹¹ In the data, low cash firms cover about 10 percent of their expenditures, so I set $x_l = 0.1$. They account for 20 percent of investment in 1980 (see Figure ??), so I use $s_k = 0.20$ as a target. The average leverage of non financial firms (debt over debt plus equity) is 0.4, as shown for instance in Graham, Leary and Roberts (forthcoming). Together with x_l and s_k , this fixes the value of $x_h = 0.62$.¹²

Beyond, x_l and x_h , the parameters governing corporate finance are the degree of moral hazard ξ and the cost of monitoring ζ . Since ζ and ξ are not really separately identified, I set a value for ξ . The restriction $\xi > x_h$ ensures that monitoring demand is positive for all firms and of course we must have $\xi < 1$ since it is measures a fraction, so I can pick any $\xi \in (0.62, 1)$. I set $\xi = 0.8$. The results do not depend on this choice. What is important, however, is the value of $\frac{\zeta}{1+\zeta}$. To identify ζ , I use a target for the unit cost of external finance $\varphi + \zeta\bar{\mu}/\bar{b}_k$ of 2.05 percent, based on the following calculation. Bonds represent about half of credit market instruments for non financial firms in 1989. For bonds, the cost is the asset management fee of 1 percent plus a liquidity premium around 0.5 percent which in equilibrium is paid to market makers (see Almeida and Philippon (2007) for a discussion) plus a 1 percent issuance fee paid every 10 year (the average bond maturity in 1989). This gives 1.6 percent. For the remaining half, I assume a cost of 2.5 percent, consistent with the spread between the reference rate and the lending rate of banks, as presented in Fixler (2009). Note that 2.5 percent is also the management fee for alternative asset managers, which presumably should be counted as monitored finance. My target for $\varphi + \zeta\bar{\mu}/\bar{b}_k$ is therefore $(1.6+2.5)/2 = 2.05$ percent.

To summarize the corporate finance calibration, I have used micro data to pin down all the parameters, except for two: the number of high-cash firms k_h , and the monitoring efficiency ζ . To make sure the model is over-identified, I attempt to match three moments: the share of low cash firms in aggregate investment $s_k = 0.20$, the unit cost of monitored finance $\varphi + \zeta\bar{\mu}/\bar{b}_k = 2.05$ percent, and the quantity of external finance over GDP $\bar{b}_k/y = 0.806$.

Household Finance In the model there is a mass 1 of young workers and a mass 1 of old workers. Aggregate labor endowment is normalized to 1, so young workers are endowed with $\frac{1}{1+\lambda}$ units of labor on average and old workers are endowed $\frac{\lambda}{1+\lambda}$ units of labor on average, and λ drives the desire of young workers to borrow.

¹¹For firms, I take into account that k/y in the model is not the same as k/y in the data because of the fluctuation in the ratio of market to book equity and because my empirical measure does not capture privately held capital. The model implies k/y of 1.7 while firm intermediation in 1989 is around 1.35. To get the correct measure of income, I therefore use $\varphi_k = \frac{1.3}{1.7} \times 1.3$ percent.

¹²External finance per unit of asset is $s_k(1 - x_l) + (1 - s_k)(1 - x_h)$. As explained, monitored finance is the stock of debt plus issuances of new debt and equity. To be internally consistent, I therefore set x_h such that $s_k(1 - x_l) + (1 - s_k)(1 - x_h) = \frac{0.806}{0.665} \times 0.4 = 0.485$. This gives $x_h = 0.619$.

Table 3: Parameters Estimated from Micro Data

High cash	Low cash	Asset Mgt Fee
$x_h = 0.62$	$x_l = 0.1$	$\varphi = 0.01$

I assume that heterogeneity η is uniform over $\left[\frac{1-H}{1+\lambda}, \frac{1+H}{1+\lambda}\right]$ for some $H > 0$. The density inside the interval is $f(\eta) = \frac{1+\lambda}{2H}$. Old workers have the same distribution, but scaled by λ . The top earners in this economy are the old workers with high values of η . More precisely, the top 40 percent of old people represent the top 20 percent of the population (this is true for the value of λ that I consider below) and have an average income of $\frac{\lambda}{1+\lambda} \left(1 + \frac{3}{5}H\right)$. Their combined labor income is $0.4 \frac{\lambda}{1+\lambda} \left(1 + \frac{3}{5}H\right) w$ relative to a total income of w . In the data the top quintile earns 46.8 percent of aggregate income in 1989 (Jones and Weinberg, 2000), but some of their income is capital income. In the model, the labor income share is α and the capital income share is $1 - \alpha$. I calibrate the model so that they earn 60 percent of capital income.¹³ Therefore, H must be such that

$$0.4 \frac{\lambda}{1+\lambda} \left(1 + \frac{3}{5}H\right) \alpha + 0.6(1 - \alpha) = 0.468. \quad (11)$$

This pins down H given λ and α .

Table 4: Estimation

	<u>Moments</u>			
	Business Debt	Unit Cost	l -Firms	HH Debt
	\bar{b}_k/y	$\varphi + \zeta \frac{\mu_k}{b_k}$	s	\bar{b}_c/y
data	0.806	0.0205	0.20	0.73
model	0.811	0.0208	0.199	0.73
	HH Particip.	Top 20 percent HH	Liquidity	Fin. Share
	$1 - F(\hat{\eta})$	Eq (11)	m/y	y^f/y
data	0.84	0.468	0.71	0.0585
model	0.84	0.468	0.71	0.0580
	<u>Implied Parameters</u>			
	Monit. Cost	H-Cash Firms	Earn. Slope	Ineq.
	$\zeta = r/3.35$	$k_h/k^* = 0.622$	$\lambda = 2.07$	$H = 0.875$
	Fix. Cost	Liq. Demand	Liq.	
	$\kappa = 0.023$	$\nu = 0.0181$	$\psi_m = 0.019$	

Calibration of model using data from 1989. k^* refers to the first best level of capital (obtained by setting all intermediation costs to zero)

I also target the participation rate of households in the credit market, and the size of the household debt market. According to Moore and Palumbo (2010), in 1989, 84 percent of heads of household of age 45 or less had positive debt balance. I therefore $1 - F(\hat{\eta}) = 0.84$, which pins down $\hat{\eta}$ given H . The size of the household debt market is $\frac{\bar{b}_c}{w} = \frac{1+\gamma}{2+r+\gamma} \int_{\eta>\hat{\eta}} \left(\left(\lambda - (1-\varphi)^{-1} \right) \eta - \kappa \right) dF(\eta)$. With a uniform distribution, we have $\int_{\eta>\hat{\eta}} \left(\left(\lambda - \frac{1}{1-\varphi} \right) \eta - \kappa \right) dF(\eta) = \frac{1+\lambda}{2H} (\bar{\eta} - \hat{\eta}) \left(\left(\lambda - \frac{1}{1-\varphi} \right) \bar{\eta} - \kappa \right)$ with $\bar{\eta} \equiv \frac{1+H}{1+\lambda}$. Aggregate household debt is therefore

$$\bar{b}_c = \frac{1+\gamma}{2+r+\gamma} \frac{1+\lambda}{2H} (\bar{\eta} - \hat{\eta}) \left(\left(\lambda - \frac{1}{1-\varphi} \right) \frac{\bar{\eta} + \hat{\eta}}{2} - \kappa \right) w. \text{ In the data, } \bar{b}_c/y \text{ is } 0.73.$$

The model is over-identified by one parameter but the fit is good. The implied parameters are all reasonable. For instance, $\kappa = 2.3$ percent means that the fixed cost is 4.3 percent of the average wage of young workers, and 10.7 percent

¹³According to Saez's data, when the top decile earns 40 percent of total income, it earns 31.5 percent of labor income. If θ is the fraction of capital income earned by this group, we have $0.315\alpha + (1 - \alpha)\theta = 0.4$ which implies that they earn therefore 60 percent of capital income. I therefore require that

of the wage of the marginal worker (the one earning $\hat{\eta}$ who is indifferent between participating or not participating). In the model, the finance industry earns 1.35 percent of GDP from liquidity, 2.08 from household credit, and 2.37 from business intermediation, for a total of 5.8 percent of GDP.

References

- Almeida, Heitor, and Thomas Philippon.** 2007. "The Risk-Adjusted Cost of Financial Distress." *Journal of Finance*, 62(6): 2557–2586.
- Altinkilic, Oya, and Robert S. Hansen.** 2000. "Are There Economies of Scale in Underwriting Fees? Evidence of Rising External Financing Costs." *Review of Financial Studies*, 13: 191–218.
- Bai, Jennie, Thomas Philippon, and Alexi Savov.** 2011. "Have Financial Markets Become More Informative?" Working Paper NYU.
- Baumol, William J.** 1967. "Macroeconomics of Unbalanced Growth: The Anatomy of the Urban Crisis." *American Economic Review*, 57: 415–426.
- Bickenbach, F, E Bode, D Dohse, A Hanley, and R Schweickert.** 2009. "Adjustment After the Crisis: Will the Financial Sector Shrink?" Kiel Policy Brief.
- Chen, Qi, Itay Goldstein, and Wei Jiang.** 2010. "Payoff complementarities and financial fragility: Evidence from mutual fund outflows." *Journal of Financial Economics*, 97(2): 239 – 262.
- DeMarzo, Peter M., and Darrell Duffie.** 1991. "Corporate financial hedging with proprietary information." *Journal of Economic Theory*, 53(2): 261–286.
- DeMarzo, Peter M., and Darrell Duffie.** 1995. "Corporate Incentives for Hedging and Hedge Accounting." *Review of Financial Studies*, 8(3): 743–71.
- Fixler, Denis.** 2009. "Price Index Concepts and Measurement." , ed. John S. Greenlees W. Erwin Diewert and Charles R. Hulten, Chapter Incorporating Financial Services in a Consumer Price Index. University of Chicago Press.
- Froot, Kenneth A., David Scharfstein, and Jeremy C. Stein.** 1993. "Risk Management: Coordinating Corporate Investment and Financing Policies." *The Journal of Finance*, 48(5): 1629–1658.
- Graham, John R, Mark Leary, and Michael R. Roberts.** forthcoming. "A Century of Capital Structure: The Leveraging of Corporate America." *Journal of Financial Economics*.
- Greenwood, Robin, and David Scharfstein.** 2013. "The Growth of Modern Finance." *Journal of Economic Perspectives*, 27(2): 3–28.
- Hood, Kyle.** 2013. "Measuring the Services of Commercial Banks in the National Income and Products Accounts." *Survey of Current Business*, 8–19.
- Jones, Arthur, and Daniel Weinberg.** 2000. "The Changing Shape of the Nation's Income Distribution: 1947-1998." US Census Bureau.
- Lucas, Robert E. Jr.** 2000. "Inflation and Welfare." *Econometrica*, 68(2): 247–274.
- Lucas, Robert E. Jr., and Nancy L. Stokey.** 1987. "Money and Interest in a Cash-in-Advance Economy." *Econometrica*, 55(3): 491–513.
- Mehra, Rajnish, Facundo Piguillem, and Edward C. Prescott.** 2011. "Costly financial intermediation in neoclassical growth theory." *Quantitative Economics*, 2(1): 1–36.
- Moore, Kevin B., and Michael G. Palumbo.** 2010. "The Finances of American Households in the Past Three Recessions: Evidence from the Survey of Consumer Finances." Staff Paper Federal Reserve Board.
- O'Mahony, Mary, and Marcel P. Timmer.** 2009. "Output, Input and Productivity Measures at the Industry Level: The EU KLEMS Database." *The Economic Journal*, 119(538): F374–F403.
- Philippon, Thomas.** 2012. "Notes on Equilibrium Financial Intermediation." mimeo NYU.
- Sargent, Thomas J., and Bruce D. Smith.** 2010. "The Timing of Tax Collections and the Structure of "Irrelevance" Theorems in a Cash-in-Advance Model." *Macroeconomic Dynamics*, 14: 585–603.
- Sidrauski, Miguel.** 1967. "Rational Choice and Patterns of Growth in a Monetary Economy." *American Economic Review*, 57: 534–544.
- Sirmans, C. F., and John D. Benjamin.** 1990. "Pricing fixed rate mortgages: Some empirical evidence." *Journal of Financial Services Research*, 4: 191–202. 10.1007/BF00365422.