

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 is 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, Piguille 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 $\left(C_t \varphi \left(\frac{M_t}{C_t} \right) \right)^{1-\rho}$. 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-Douglass 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 investments 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 $\Xi_t \equiv (1 + \nu)(\log A_t + \beta \log A_{t+1}) + \nu \log \frac{\nu}{\psi_{m,t}} + \beta \nu \log \frac{\nu}{\psi_{m,t+1}}$ 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

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

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

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