On Fintech and Financial Inclusion

Thomas Philippon*

June 2019

Abstract

This paper makes three points. I first update previous work to show that the cost of financial intermediation has declined since the Great Recession, thanks to technology and increased competition. I then analyze two features of new financial technologies: returns to scale and big data. I argue that changes in the nature of fixed versus variable costs are likely to improve access to financial services and reduce inequality. Big data, on the other hand, can generate a trade-off between efficiency and discrimination.

JEL: E2, G2, N2

*Stern School of Business, New York University; NBER and CEPR. This paper was prepared for the 2019 Annual Conference of the BIS. I am grateful to my discussants [..] and to Marcos Sonnervig for outstanding research assistance.

FinTech covers digital innovations and technology-enabled business model innovations in the financial sector. Such innovations can disrupt existing industry structures and blur industry boundaries, facilitate strategic disintermediation, revolutionize how existing firms create and deliver products and services, provide new gateways for entrepreneurship, democratize access to financial services, but also create significant privacy, regulatory and law-enforcement challenges. Examples of innovations that are central to FinTech today include cryptocurrencies and the blockchain, new digital advisory and trading systems, artificial intelligence and machine learning, peer-to-peer lending, equity crowdfunding and mobile payment systems.

I update the work of Philippon (2015) with post-crisis U.S. data. I find that the unit cost of financial intermediation has begun to decline over the past 5 years.

I then study two issues that are at the heart of the fintech debate: access to finance and discrimination. If we accept the fact that fintech brings efficiency gains to financial intermediation, the next question is: how will these gains be shared? Will fintech democratize access to financial services or will it increase inequality?

I highlight two effects that will shape the answer to these questions. The first is the impact of returns to scale brought by technology. I show that the nature of fixed versus variable costs has changed and I argue that it is likely to reduce inequality. On the other hand, I argue that big data can lead to more discrimination.

Recent literature Philippon (2016) discusses the literature up to 2016 so I will mention here only recent papers. Focusing on residential mortgages, Buchak et al. (2018) study the growth in the market share of shadow bank and FinTech lenders, arguing that it can be explained by differences in regulation and technological advantages. They find that FinTech lenders serve more creditworthy borrowers (relative to shadow banks) and charge higher interest rates (14-16 basis points), what is consistent with the evidence that financial technology have failed to reduce intermediation costs (Philippon, 2015).

The higher interest paid to FinTech lenders may also be a result of the convenience provided by them. Fuster et al. (2019) study the differences between FinTech and traditional lenders in the mortgage market and find that the former is quicker in processing applications (20% faster), without increasing loan risk. They also provide evidence that FinTech lenders adjust supply more elastically to demand shocks and increase the propensity to refinance, especially among borrowers that are likely to benefit from it. Their results suggest that FinTechs have improved the efficiency of financial intermediation in mortgage markets.

The advent of FinTech is often seen as a promising avenue for reducing inequality in access to credit. Bartlett et al. (2018) study this issue, analyzing the role of FinTech lenders in alleviating discrimination in mortgage markets. They find that, FinTechs and face-to-face lenders are equally discriminatory in charging higher interest rates from minorities. However, in the loan accept/reject decision (as opposed to pricing), FinTechs appear to be much less discriminatory.

Regarding the use of new technologies in credit markets, Berg et al. (2019) analyses the information content of the "digital footprint" (an easily accessible information for any firm conducting business in the digital sphere) for predicting consumer default. With data from a German E-commerce, they find that it equals or exceeds the predictive power of traditional credit bureau scores. Their results suggest that FinTech lenders might have a superior ability for screening borrowers.

FinTechs are also competing with traditional financial institutions in the market for wealth management. While, currently, in most of the industry this is done through personalized individual assistance, FinTechs provide it through automated financial advisors, also known as "robo-advisors". The United States is, by far, the leading market for robo-advisors. In 2017, it had more robo-advisors than any other economy in the world (about 200), with 57 percent of all investments in robo-advisors (Abraham et al., 2019). Nevertheless, the amount of assets managed by robo-advisors is still a small portion of total assets under management, with average client wealth much smaller than the average in the industry (Economist, 2017).

Abraham et al. (2019) argues that because they save on fixed costs (such as salaries of financial advisors or maintenance of physical offices), robo-advisors can reduce minimum investment requirements and charge lower fees.

1 Inefficiency of the Existing System

The main finding in Philippon (2015) is that the unit cost of financial intermediation in the U.S. has remained around 2% for the past 130 years. Improvements in information technologies have not been passed through to the end users of financial services. This section offers an update of this work.

1.1 Financial Expenses and Intermediated Assets

To organize the discussion I use a simple model economy consisting of households, a non-financial business sector, and a financial intermediation sector. The details of the model are in the Appendix. The income share of finance, shown in Figure 1, is defined as[1]

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

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

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

[1]Philippon (2015) discusses various issues of measurement. 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 U.S. 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 U.S. economy. Philippon (2015) also explains the robustness of the main findings to large changes in government spending (because of wars), the rise of services (finance as a share of services displays a similar pattern to the one presented here), globalization (netting out imports and exports of financial services).

where $b_{c,t}$ is consumer credit, m_t are assets providing liquidity services, 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 historical evidence.[2]

Measuring intermediated assets is complicated because these assets are heterogenous. As far as corporate finance is concerned, the model is fundamentally a user cost model. Improvements in corporate finance (a decrease in ψ_k) lower the user cost of capital and increase the capital stock, which, from a theoretical perspective, should include all intangible investments and should be measured at market value. 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. The model also incorporates liquidity services provided by specific liabilities (deposits, checking accounts, some form of repurchase agreements) issued by financial intermediaries. One can always write the RHS of (1) as $\psi_{c,t}\left(b_{c,t} + \frac{\psi_{m,t}}{\psi_{c,t}} m_t + \frac{\psi_{k,t}}{\psi_{c,t}} k_t\right)$. Philippon (2015) finds that the ratios $\frac{\psi_{m,t}}{\psi_{c,t}}$ and $\frac{\psi_{k,t}}{\psi_{c,t}}$ are close to one.[3] As a result one can define intermediated assets as

$$q_t \equiv b_{c,t} + m_t + k_t. \tag{2}$$

The principle is to measure the instruments on the balance sheets of non-financial users, households and non-financial firms. This is the correct way to do the accounting, rather than looking at the balance sheet of financial intermediaries. 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 for the non-financial sector, displayed in Figure 1.

1.2 Unit Cost and Quality Adjustments

I can then divide the income of the finance industry by the quantity of intermediated assets to obtain a measure of the unit cost

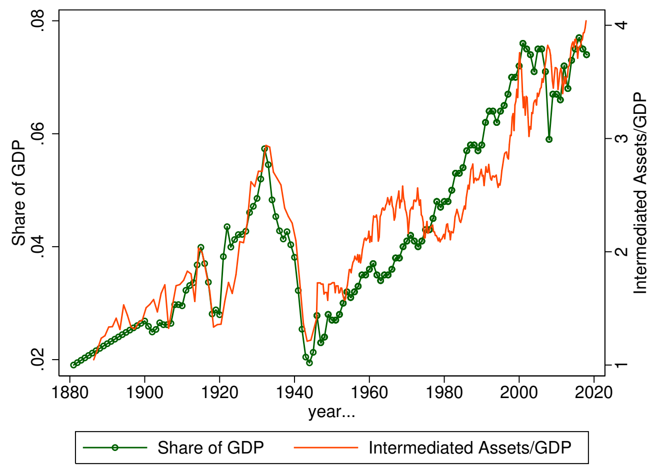
$$\psi_t \equiv \frac{y_t^f}{q_t}. \tag{3}$$

Figure 2 shows that this unit cost is around 2% and relatively stable over time. In other words, I estimate that it costs two cents per year to create and maintain one dollar of intermediated financial asset. Equivalently, the annual rate of return of savers is on average 2 percentage points below the funding cost of borrowers. The updated series

[2] The fact that the finance share of GDP 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% in 1980, 4.68% in 1990, 4.19% in 2000, and 4.47% in 2006.

[3] This is true most of the time, but not when quality adjustments are too large. Philippon (2015) provides calibrated quality adjustments for the U.S. financial system.

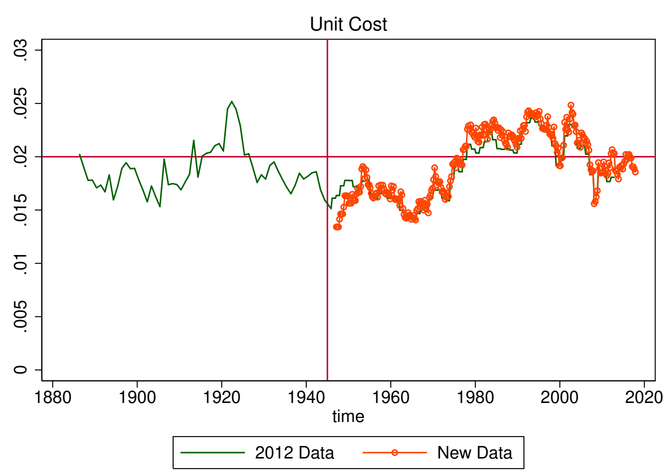
Figure 1: Finance Income and Intermediated Assets



Notes: Both series are expressed as a share of GDP. Finance Income is the domestic income of the finance and insurance industries, i.e., aggregate income minus net exports. Intermediated Assets include debt and equity issued by non financial firms, household debt, and various assets providing liquidity services. Data range for Intermediated Assets is 1886 - 2012. See Philippon (2015) for historical sources and details about the underlying data.

are similar to the ones in the original paper. The unit costs for other countries are estimated by Bazot (2013) who finds convergence to US levels.

Figure 2: Unit Cost of Financial Intermediation

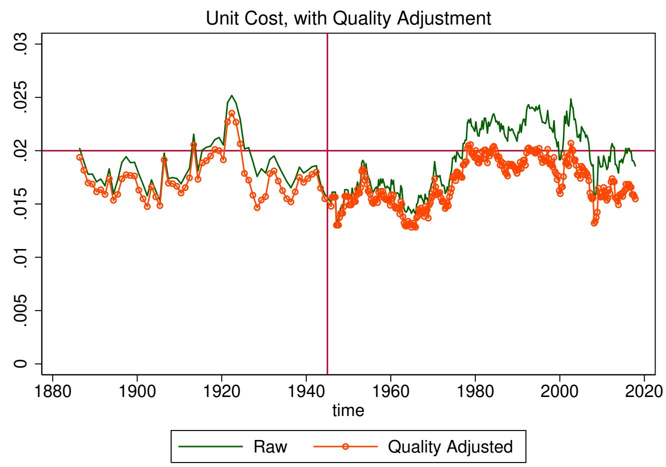


Notes: The raw measure is the ratio of finance income to intermediated assets, displayed in Figure 1. The 2012 data is from Philippon (2015), while the new data was accessed in May 2016. Data range is 1886 - 2015.

The raw measure of Figure 2, however, does not take into account changes in the characteristics of borrowers. These changes require quality adjustments to the raw measure of intermediated assets. For instance, 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. Similarly, with household finance, 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.[4] Measurement problems arise when the mix of high- and low-quality borrowers changes over time.

Following Philippon (2015), I then perform a quality adjustment to the intermediated assets series. Figure 3 shows the quality adjusted unit cost series. It is lower than the unadjusted series by construction since quality adjusted assets are (weakly) larger than raw intermediated assets. The gap between the two series grows when there is entry of new firms, and/or when there is credit expansion at the extensive margin (i.e., new borrowers). Even with the adjusted series, however, we do not see a significant decrease in the unit cost of intermediation over time.

Figure 3: Unit Cost and Quality Adjustment



Notes: The quality adjusted measure takes into account changes in firms' and households' characteristics. Data range is 1886 - 2015.

As I have argued in the past, the puzzle is why we have not seen substantial productivity gains in financial intermediation. The good news is that, however late, these gains might be happening now.

2 A Tale of Two Fixed Costs

We consider a simple model of imperfect competition in asset management services. We emphasize the role of technology, and fixed costs in particular. The key point is that there are two types of fixed costs: fixed costs to set up a business or a system or a platform; and then fixed cost per relationship with each client.

[4]Using 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%. 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% between 1989 and 2007. In the mortgage market, Mayer and Pence (2008) show that subprime originations account for 15% to 20% of all HMDA originations in 2005.

2.1 Setup

There is a collection of identical financial intermediaries and a continuum of mass 1 of households whose wealth is distributed according to the (cumulative) distribution G. In period 1, financial intermediaries decide whether to enter the market or not. There is a fixed entry cost equal to Φ. Investment decisions are made in period 2. Households consume in period 3. We assume that they are risk neutral, or we interpret the returns as being risk adjusted.

In period 2, households decide whether to invest their wealth by themselves or with the financial intermediaries. If they invest by themselves ("directly"), they get a rate of return $\underline{R}$, at no cost. If they invest through the intermediaries, they get a return $\bar{R}$, and pay a fee $f(w_i)$, which depends on their wealth w_i. Financial intermediaries have a fixed cost per client equal to ϕ and access to a technology that delivers return $\bar{R}$. We restrict attention to linear fees where the baseline covers the fixed cost, of the form

$$f\left(w\right) = \phi + f_1(N)w,$$

where $f_1(.)$ is a decreasing function and N is the number of financial intermediaries that entered the asset management market in period 1. This specification captures various forms of competition, including constant markups. We restrict f_1 to be weakly positive, decreasing and lower that the monopoly price.

The model can be solved by backward induction. In period 2, households optimal decision between intermediation or direct investment is described by the cutoff rule:

$$U\left(inter, w_i\right) > U\left(direct, w_i\right)$$

This can be written as $w_i\bar{R} - f\left(w_i\right) > w_i\underline{R}$ and it leads to as simple cutoff rule. Investors poorer than $\bar{w}$ do not use asset management services, where we define

$$\bar{w}\left(N\right) \equiv \frac{\phi}{\bar{R} - \underline{R} - f_1\left(N\right)}. \tag{4}$$

A fraction $G\left(\bar{w}\right)$ will therefore invest directly, and a fraction $1 - G\left(\bar{w}\right)$ will use intermediation services. We consider a symmetric equilibrium where intermediaries have the same number of clients. The net profits of one intermediary is therefore

$$\pi\left(N\right) \equiv \frac{f_1\left(N\right)}{N} \int_{\bar{w}(N)}^{\infty} w dG\left(w\right)$$

where $\bar{w}\left(N\right)$ is defined in equation (4). Note that N should be interpreted as the number of asset manager per capita since we have normalized population to one.

We next consider the entry decision in period 1. Free entry requires

$$\pi(N) \geq \Phi$$

with equality if entry is positive. This pins down the number of firms entering the market and the equilibrium markup f_1. Finally, welfare in this economy is characterized by the following expression:

$$\begin{aligned} W &= \int_0^{\bar{w}} \underline{R} w dG(w) + \int_{\bar{w}}^{\infty} \left(\bar{R}w - f(w)\right) dG(w) + N\pi - N\Phi \\ &= \int_0^{\bar{w}} \underline{R} w dG(w) + \int_{\bar{w}}^{\infty} \left(\bar{R}w - \phi\right) dG(w) - N\Phi \end{aligned}$$

Definition 1. Given the cost structure (Φ, ϕ), an equilibrium is a number of asset managers N such that $\pi(N) = \Phi$.

2.2 The impact of fintech

Now we can analyze equilibria under alternative cost structures. We define the traditional "Banking" or "traditional asset management" equilibrium $(\bar{w}^B, N^B)$, as the equilibrium associated to the cost structure (Φ^L, ϕ^H); and the "Fintech equilibrium" $(\bar{w}^F, N^F)$, as the equilibrium associated to the cost profile (Φ^H, ϕ^L). We assume:

$$\Phi^H > \Phi^L \text{ and } \phi^H > \phi^L$$

The key idea is that fintech requires an upfront investment but then it has a lower cost per relationship. The best example is a robo advisor. It is costly to program the robo $\left(\Phi^H > \Phi^L\right)$. But then it does not matter if the robo has 100 clients or 1 million clients $\left(\phi^H > \phi^L\right)$ and perhaps even $\phi^L = 0$. The literature so far has failed to capture the difference between these two types of fixed costs. I will show that they have very different welfare implications.

The following proposition establishes that if in the Banking equilibrium, a FinTech can make profit by charging the same fee as the established intermediaries $\left(f(w) = \phi^H + f_1(N^B)w\right)$ then there are more households using asset management services in the FinTech equilibrium $(\bar{w}^F < \bar{w}^B)$.

Proposition 1. *If* $\Phi^H - \Phi^L \leq \left(\phi^H - \phi^L\right) \frac{1-G\left(\bar{w}^B\right)}{N^B}$, *then the wealth cutoff is smaller in the FinTech equilibrium* $(\bar{w}^F < \bar{w}^B)$.

Proof. There are two cases to consider, depending on the relative number of intermediaries in the two equilibrium.

CASE 1: $N^F \geq N^B$. Recall that f_1 is decreasing in N. Hence the result follows immediately from the definition of $\bar{w}$:

$$\bar{w}^F = \frac{\phi^L}{\bar{R} - \underline{R} - f_1(N^F)} \leq \frac{\phi^H}{\bar{R} - \underline{R} - f_1(N^B)} = \bar{w}^B$$

CASE 2: $N^F < N^B$. First note that the profit of a FinTech intermediary entering the banking equilibrium and charging $f(w) = \phi^H + f_1(N^B)w$ is:

$$\pi(N^B; \{f(w) = \phi^H + f_1(N^B)w\}; \phi^L) = \frac{1}{N^B}\int_{\bar{w}^B}^{\infty}\left[\phi^H + f_1(N^B)w - \phi^L\right]dG(w) = \underbrace{\frac{1}{N^B}f_1(N^B)\int_{\bar{w}^B}^{\infty} wdG(w)}_{\Phi^L} + \frac{[\phi^H - \phi^L]\mathbb{P}(w > \bar{w}^B)}{N^B}$$

Now define: $\bar{\Phi} \equiv \frac{1}{N^F}\left(\bar{R} - \underline{R} - \frac{\phi^L}{\bar{w}^B}\right)\int_{\bar{w}^B}^{\infty} wdG(w)$ and note that $\bar{\Phi} > \Phi^H$:

$$\begin{aligned}
\bar{\Phi} &= \frac{1}{N^F}\left\{\underbrace{\left(\bar{R} - \underline{R} - \frac{\phi^H}{\bar{w}^B}\right)}_{f_1(N^B)} + \frac{[\phi^H - \phi^L]}{\bar{w}^B}\right\}\int_{\bar{w}^B}^{\infty} wdG(w) = \\
&= \frac{N^B}{N^F}\left\{\Phi^L + [\phi^H - \phi^L]\frac{\int_{\bar{w}^B}^{\infty} wdG(w)}{N^B\bar{w}^B}\right\} \geq \\
&\geq \frac{N^B}{N^F}\left\{\Phi^L + [\phi^H - \phi^L]\frac{\int_{\bar{w}^B}^{\infty} wdG(w)}{N^B\mathbb{E}\left[w|w > \bar{w}^B\right]}\right\} = \\
&= \frac{N^B}{N^F}\left\{\Phi^L + [\phi^H - \phi^L]\frac{\mathbb{P}(w > \bar{w}^B)}{N^B}\right\} \geq \\
&\geq \frac{N^B}{N^F}\Phi^H > \Phi^H,
\end{aligned}$$

where the second equality follows from the free entry condition in the Banking equilibrium, the first inequality follows from the fact that $\bar{w}^B \leq \mathbb{E}\left[w|w > \bar{w}^B\right]$ and the second inequality reflects the assumption. Now, by free entry in the FinTech equilibrium we have $\Phi^H = \pi(N^F) = \frac{1}{N^F}\left(\bar{R} - \underline{R} - \frac{\phi^L}{\bar{w}^F}\right)\int_{\bar{w}^F}^{\infty} wdG(w)$. And finally, because $\left(\bar{R} - \underline{R} - \frac{\phi^L}{\bar{w}}\right)\int_{\bar{w}}^{\infty} wdG(w)$ is increasing in $\bar{w}$, $\bar{\Phi} > \Phi^H \implies \bar{w}^B \geq \bar{w}^F$, establishing the result. □

The key take-away from this analysis is that Fintech leads to lower inequality in access to services. By lowering the fixed cost per relationship is allows more households to benefit from advisory services. Moreover, in equilibrium, the rich subsidize the poor. The rich pay the lion's share of fees that serve to cover the fixed cost Φ. Once this cost is paid, poor households benefit from cheap services.

3 Discrimination

Let us now discuss a potentially darker side of fintech. Big data can increase discrimination. We present the idea in the context of a competitive lending market. We make it competitive for simplicity and because we have already discussed imperfect competition above.

3.1 Setup

We assume that lenders offer the breakeven rate R, which is a function of an unknown borrower characteristic x, which can be interpreted as default risk.

$$R_{\mathbb{I}} = \mathbb{E}[x|\mathbb{I}],$$

where $\mathbb{I}$is the information set of the lender. Households are divided in two groups $z = A, B$: minorities ($z = A$; measure p) and non-minorities ($z = B$, measure $1 - p$). As usual in the literature of statistical discrimination (Aigner and Cain, 1977), we assume that the distribution of x is different between the two groups. In particular we assume that it is normally distributed, with different means

$$x_{min} \sim N(a, \sigma^2)$$

and

$$x_{non-min} \sim N(b, \sigma^2)$$

with $a > b$. Thus in our model minorities are riskier.

There are two types of lenders: FinTechs and banks. We assume that banks observe a signal about x:

$$\tilde{x}^B = x + \epsilon_1$$

where $\epsilon_1 \sim N(0, \sigma_B^2)$. FinTech lenders, on the other hand, observe a signal and also a "proxy" for the minority dummy (z):

1. Signal: $\tilde{x}^F = x + \epsilon_{1f} \quad , \epsilon_{1f} \sim N(0, \sigma_F^2), \quad \sigma_F^2 \leq \sigma_B^2$

2. Proxy : $y = z + \epsilon_2 \quad , \epsilon_2 \sim N(0, \sigma_y^2)$

The idea is that fintech have access to two types of data. Standard data similar to that used by banks. And new data from social media or other sources. These data provide information about x but also contain proxies for z. As a result, using these data improves ϵ_1 so that $\sigma_F^2 \leq \sigma_B^2$, but it also reveal "too much" information about z.

Lenders solve a signal extraction problem. The solution for traditional lenders is:

$$\mathbb{E}[x|\tilde{x}] = (1 - \beta)\bar{x} + \beta\tilde{x},$$

where

$$\beta = \frac{\sigma^2 + p(1-p)(a-b)^2}{\sigma^2 + p(1-p)(a-b)^2 + \sigma_B^2},$$

and $\bar{x} = pa + (1-p)b$ is the unconditional mean of x.

For the FinTech lenders, the solution is:

$$\mathbb{E}[x|\tilde{x}, y] = (1-\gamma_1-\gamma_2)\bar{x} + \gamma_1[ya + (1-y)b] + \gamma_2\tilde{x},$$

where

$$\gamma_1 = \frac{\sigma_F^2\left[(a-b)^2 p(1-p)\right]}{\left\{(a-b)^2\left[p(1-p)+\sigma_y^2\right]\right\}\left[\sigma^2+\sigma_F^2+(a-b)^2p(1-p)\right] - \left[(a-b)^2p(1-p)\right]^2}$$

and

$$\gamma_2 = \frac{\left\{(a-b)^2\left[p(1-p)+\sigma_y^2\right]\right\}\left[\sigma^2+(a-b)^2p(1-p)\right] - \left[(a-b)^2p(1-p)\right]^2}{\left\{(a-b)^2\left[p(1-p)+\sigma_y^2\right]\right\}\left[\sigma^2+\sigma_F^2+(a-b)^2p(1-p)\right] - \left[(a-b)^2p(1-p)\right]^2}$$

Note that when the proxy for "minority group" is perfect ($\sigma_y^2 = 0$), then:

$$\gamma_1 = \frac{\sigma_F^2}{\sigma^2+\sigma_F^2}$$

$$\gamma_2 = \frac{\sigma^2}{\sigma^2+\sigma_F^2}$$

3.2 Planner

The planner faces a trade-off between efficiency and discrimination. We define discrimination as the difference between the rates offered to members of different groups with the same "productive characteristic" x. Also, we use the mean square error as as measure of inefficiency. The planner therefore minimizes the loss function:

$$L(\mathbb{I}) = \underbrace{\mathbb{E}\left[\left(x-R_{\mathbb{I}}\right)^2\right]}_{\text{Inefficiency}} - \theta\underbrace{\left\{\mathbb{E}\left[R_{\mathbb{I}}\middle|x,\text{minority}\right] - \mathbb{E}\left[R_{\mathbb{I}}\middle|x,\text{non-minority}\right]\right\}}_{\text{Discrimination}}$$

3.2.1 Benchmark: "Minority observed"

When "minority" and signal $\tilde{x}$ are included in the information set $\mathbb{I}$:

$$\text{Discrimination} = \mathbb{E}\left[R_{\mathbb{I}}\middle|x,\text{minority}\right] - \mathbb{E}\left[R_{\mathbb{I}}\middle|x,\text{non-minority}\right] = \frac{\sigma_B^2}{\sigma^2+\sigma_B^2}[a-b]$$

$$\text{Inefficiency} = \mathbb{E}\left[\left(x-R_{\mathbb{I}}\right)^2\right] = \left(\frac{\sigma_B^2}{\sigma^2+\sigma_B^2}\right)^2\sigma^2 + \left(\frac{\sigma^2}{\sigma^2+\sigma_B^2}\right)^2\sigma_B^2$$

3.2.2 Banks

When only the signal $\tilde{x}$ is included in the information set $\mathbb{I}$:

$$\text{Discrimination} = \mathbb{E}\left[R_{\mathbb{I}} \middle| x, \text{minority}\right] - \mathbb{E}\left[R_{\mathbb{I}} \middle| x, \text{non-minority}\right] = 0$$

$$\text{Inefficiency} = \mathbb{E}\left[\left(x - R_{\mathbb{I}}\right)^2\right] = (1-\beta)^2\left[(a-b)^2 p(1-p) + \sigma^2\right] + \beta^2\sigma_B^2$$

Where:

$$\beta = \frac{\sigma^2 + p(1-p)(a-b)^2}{\sigma^2 + p(1-p)(a-b)^2 + \sigma_B^2}$$

3.2.3 FinTech

When the proxy and signal $\tilde{x}$ are included in the information set $\mathbb{I}$:

$$\text{Discrimination} = \mathbb{E}\left[R_{\mathbb{I}} \middle| x, \text{minority}\right] - \mathbb{E}\left[R_{\mathbb{I}} \middle| x, \text{non-minority}\right] = \gamma_1[a-b]$$

$$\text{Inefficiency} = \mathbb{E}\left[\left(x - R_{\mathbb{I}}\right)^2\right] = (1-\gamma_1-\gamma_2)^2\left[(a-b)^2 p(1-p) + \sigma^2\right] + \gamma_1^2\left[(a-b)^2\sigma_y^2 + \sigma^2\right] + \gamma_2^2\sigma_{\tilde{x}}^2$$

3.3 Discussion

FinTech lenders may differ from banks in two dimensions: the precision of their signal $\tilde{x}^F$ and the precision of the proxy y. If the proxy is uninformative, then there is no trade-off between inefficiency and discrimination: FinTechs would achieve a better outcome on both dimensions. If, on the other hand, banks and FinTechs' signal have the same variance ($\sigma_B^2 = \sigma_F^2$), then efficiency is achieved at the cost of higher discrimination. In this case, the net welfare gain depends on the planner's weight θ.

4 Conclusion

Even in the best case scenario, FinTech is likely to create new issues. [TBC]

Appendix

A A Simple Model of Financial Intermediation Accounting

In this Appendix I sketch a model, based on Philippon (2015), that can be used for financial intermediation accounting. 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.[5] I consider a setup with two types of households: some households are infinitely lived, the others belong to an overlapping generations structure.[6] Households in the model do not lend directly to one another. They lend to intermediaries, and intermediaries lend to firms and to other households.

A.1 Technology and Preferences

Long-Lived Households

Long-lived households (index l) are pure savers. They own the capital stock and have no labor endowment. Liquidity services are modeled as money in the utility function. The households choose consumption C and holdings of liquid assets M to maximize

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

I specify the utility function as $u\left(C_t, M_t\right) = \frac{\left(C_t M_t^{\nu}\right)^{1-\rho}-1}{1-\rho}$. 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

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

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

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

The liquidity demand equation $u_M\left(t\right) = \psi_{m,t} u_C\left(t\right)$ 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 on 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 $u\left(C_{1,t}, M_{1,t}\right) + \beta u\left(C_{2,t+1}, M_{2,t+1}\right)$. 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 $C_{1,t} + \psi_{m,t} M_{1,t} = \eta_1 W_{1,t} + \left(1-\psi_{c,t}\right) B_t^c$. The screening and monitoring cost is $\psi_{c,t}$ per unit of borrowing. In the second period, the household consumes $C_{2,t+1} + \psi_{m,t+1} M_{2,t+1} = \eta_2 W_{t+1} - \left(1+r_{t+1}\right) B_t^c$. The Euler equation for OLG households is

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

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

[5] Heterogeneity and quality adjustments are discussed in Philippon (2015).

[6] The 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.

Non Financial Businesses

Non-financial output is produced under constant returns technology, and for simplicity I assume that the production function is Cobb-Douglass:[7]

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

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. 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:

$$(1-\alpha)\left(\frac{A_t n_t}{K_t},\right)^{\alpha} = r_t + \delta + \psi_{k,t}. \tag{5}$$

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

$$\alpha\left(\frac{A_t n_t}{K_t},\right)^{\alpha-1} = \frac{W_t}{A_t}. \tag{6}$$

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 of financial intermediaries is then

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

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

A.2 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

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

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

$$\beta(1+r) = (1+\gamma)^{\theta}. \tag{7}$$

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 (5) becomes

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

[7] Philippon (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-Douglass technology. Qualitatively different results only happen for elasticity values above 6, which is far above the range of empirical estimates. Thus assuming a Cobb-Douglass technology does not entail much loss of generality.

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 (7), the Euler equation of short lived households is simply

$$c_1 = (1-\psi_c)^{\frac{1}{\theta}} c_2. \tag{8}$$

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

$$\frac{b_c}{w} = \frac{(1-\psi_c)^{\frac{1}{\theta}} \eta_2 - \eta_1}{1-\psi_c + (1-\psi_c)^{\frac{1}{\theta}} \frac{1+r}{1+\gamma}}. \tag{9}$$

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

$$c = \frac{1}{1+\nu} \left(w - \psi_c b_c + (r-\gamma) k \right). \tag{10}$$

The comparative statics are straightforward. The ratios are constant along a balanced growth path with constant intermediation technology, constant demographics, and constant firms' characteristics. Improvements in corporate finance increase y, w, k/y, c/y and m/y, but leave b^c/y constant. Improvements in household finance increase b^c/y, c/y and m/y, but do not affect k. Increases in the demand for intermediation increase the finance income share ϕ while supply shifts have an ambiguous impact.

The utility flow at time t is $u(c,m) = \frac{(cm^{\nu})^{1-\rho}}{1-\rho}$ and since $m = \frac{\nu c}{\psi_m}$, we have

$$u(c,m) = \frac{\left(\frac{\nu}{\psi_m}\right)^{\nu(1-\rho)} c^{(1+\nu)(1-\rho)} - 1}{1-\rho}$$

Imagine $A = 1$ for simplicity. Then welfare for a particular generation is

$$\begin{aligned} W &= u(c_1, m_1) + \beta u(c_2, m_2) + \frac{\omega}{1-\beta} u(c_l, m_l) \\ &= \frac{\left(\frac{\nu}{\psi_m}\right)^{\nu(1-\rho)}}{1-\rho} \left(c_1^{1-\theta} + \beta c_2^{1-\theta} + \omega \frac{c_l^{1-\theta}}{1-\beta} \right) - \frac{1}{1-\rho} \end{aligned}$$

where ω is the Pareto weight on the long lived agents.

A.3 On fixed costs

The following proposition establishes that if, in the Banking equilibrium, a FinTech cannot make profit by charging a fee equal to $f(w) = \phi^L + f_1(N^B) w$ then, in the FinTech equilibrium there are fewer firms operating. For what follows it is useful to define the profit as a function of N and the intermediation cost ϕ:

$$\pi(N,\phi) = \frac{1}{N} f_1(N) \int_{\left[\frac{\phi}{\Delta - f_1(N)}\right]}^{\infty} w dG(w)$$

Proposition 2. *If $\Phi^H \geq \pi(N^B, \phi^L)$, then there is more concentration in the FinTech equilibrium ($N^F < N^B$)*

Proof. By free entry condition in the FinTech equilibrium: $\Phi^H = \pi(N^F, \phi^L)$. The inequality is then:

$$\pi(N^F, \phi^L) \geq \pi(N^B, \phi^L)$$

Because $\pi(., \phi^L)$ is decreasing in N, the inequality implies $N^F \leq N^B$. □

References

Abraham, F., S. L. Schmukler, and J. Tessada (2019). Robo-advisors: Investing through machines. *World Bank Policy Research Working Paper* (134881).

Aigner, D. J. and G. G. Cain (1977). Statistical theories of discrimination in labor markets. *ILR Review 30*(2), 175–187.

Bartlett, R., A. Morse, R. Stanton, and N. Wallace (2018). Consumer-lending discrimination in the era of fintech. Working paper.

Bazot, G. (2013). Financial consumption and the cost of finance: Measuring financial efficiency in europe (1950-2007). Working Paper Paris School of Economics.

Berg, T., V. Burg, A. Gombović, and M. Puri (2019). On the rise of fintechs – credit scoring using digital footprints. Working paper.

Bickenbach, F., E. Bode, D. Dohse, A. Hanley, and R. Schweickert (2009, October). Adjustment after the crisis: Will the financial sector shrink? Kiel Policy Brief.

Buchak, G., G. Matvos, T. Piskorski, and A. Seru (2018). Fintech, regulatory arbitrage, and the rise of shadow banks. *Journal of Financial Economics 130*(3), 453 – 483.

Economist, T. (2017). Silicon speculators.

Fuster, A., M. Plosser, P. Schnabl, and J. Vickery (2019). The role of technology in mortgage lending. *The Review of Financial Studies 32*(5), 1854–1899.

Lucas, R. E. J. (2000, March). Inflation and welfare. *Econometrica 68*(2), 247–274.

Mayer, C. and K. Pence (2008). Subprime mortgages: What, where, and to whom? Staff Paper Federal Reserve Board.

Moore, K. B. and M. G. Palumbo (2010, June). The finances of american households in the past three recessions: Evidence from the survey of consumer finances. Staff Paper Federal Reserve Board.

O'Mahony, M. and M. P. Timmer (2009). Output, input and productivity measures at the industry level: The eu klems database. *The Economic Journal 119*(538), F374–F403.

Philippon, T. (2012). Equilibrium financial intermediation. Working Paper NYU.

Philippon, T. (2015). Has the us finance industry become less efficient? on the theory and measurement of financial intermediation. *The American Economic Review 105*(4), 1408–38.

Philippon, T. (2016). The fintech opportunity.